

# **Stochastic Processes in Finance**

May 6, 2024

P.Jacobs (M&MR/RMO)

# Contents

<b>1 Motivation</b>	<b>3</b>
1.1 Risk Modelling . . . . .	3
1.2 Derivative Valuation . . . . .	3
<b>2 Stochastic Calculus</b>	<b>5</b>
2.1 Basic Definitions . . . . .	5
2.1.1 Random Variables . . . . .	5
2.1.2 Stochastic Processes . . . . .	10
2.2 Brownian Motion . . . . .	13
2.3 Itô integral with respect to Brownian motion . . . . .	14
2.4 Itô's Lemma . . . . .	19
2.5 Itô processes . . . . .	22
2.6 Semimartingales . . . . .	23
<b>3 Derivative Valuation</b>	<b>24</b>
3.1 Fundamental Asset Pricing Theorems . . . . .	24
3.2 Partial Differential Equation for a claim . . . . .	27
3.3 Partial differential equation for the density . . . . .	28
3.4 Girsanov's Theorem . . . . .	29
3.5 Derivative pricing in multi-curve paradigm . . . . .	33
<b>4 Basic models</b>	<b>35</b>
4.1 Geometric Brownian motion . . . . .	35
4.2 Vasicek's short rate model . . . . .	38
4.3 Hull-White's model . . . . .	42
4.4 CIR model . . . . .	46
4.5 CIR++ model . . . . .	48
<b>5 Interest rate market models</b>	<b>50</b>
5.1 Caplets . . . . .	50
5.2 LIBOR market model . . . . .	51
5.3 Swap Market Model . . . . .	52
5.4 Multi-curve paradigm in interest rate modeling . . . . .	53

<b>6 Heston's Model</b>	<b>57</b>
6.1 Volatility Smile . . . . .	57
6.2 Heston's Model . . . . .	58
<b>7 Lévy Processes</b>	<b>62</b>
7.1 Poisson Process . . . . .	62
7.2 Compound Poisson process . . . . .	63
7.3 Lévy Process . . . . .	66
7.4 Gamma process . . . . .	68
7.5 Variance Gamma process . . . . .	72
<b>8 Time Changed Lévy Processes</b>	<b>73</b>
8.1 Chronometers . . . . .	73
8.2 Pricing plain vanilla calls via characteristic functions . . . . .	74
<b>9 Appendices</b>	<b>77</b>
9.1 Density function of the CIR short rate . . . . .	77
9.2 Solving the Feynman-Kac equation for zero coupon bond prices in a CIR model . . . . .	80

# Chapter 1

## Motivation

### 1.1 Risk Modelling

In each of the following applications it is necessary to consider a large number of scenarios for some assets, not just at one future date but over a whole range of dates. We thus need a stochastic process simulating possible paths that may occur in the real world. The probabilities of occurrence of subsets of paths should correspond to the observed or expected distribution in the real world. Such scenarios (together with the probabilities) are sometimes called real world scenarios.

- Counterparty Credit Risk for OTC transactions: How much money can we lose on average if a counterparty on a portfolio of OTC contracts defaults? What amount can be lost in extreme scenarios?
- Value at risk: How much money can we lose on a portfolio of financial instruments over a given holding period within a given confidence level?
- Constant Portfolio Protection Instruments: How does the protection mechanism work in certain scenarios?
- Hedging: Test how well a hedge portfolio mimics the payoff of a real portfolio in a large number of scenarios?

### 1.2 Derivative Valuation

The use of stochastic processes in valuating instruments is based on the so-called fundamental asset pricing theorem. It states that we can compute the fair value of a derivative on an asset by computing its discounted expected payoff in the so-called risk neutral world, meaning that we need to change the probabilities of the stochastic payoff in a specific way. In such a risk neutral world, the fair value of the asset itself must also equal its discounted expected value at any future

time. Often the model for the asset is specified directly in the risk neutral world and calibrated such that this condition holds. In the risk neutral world, stock prices will have an expected growth rate equal to the risk free spot rate.

Note that the payoff of more complex derivatives (such as lookback options) may depend on the value of the asset at different future times. Hence we need a stochastic process to model the asset, not just a random variable.

# Chapter 2

# Stochastic Calculus

We refer to [Mik98] for a nice and intuitive introduction to stochastic calculus. The stochastic calculus lectures [Bau] provide a nice overview of the most important theorems. A thorough more general treatment of the theory of stochastic calculus can be found in [JS03].

## 2.1 Basic Definitions

### 2.1.1 Random Variables

**$\sigma$ -algebra:** A  $\sigma$ -algebra  $\mathcal{F}$  on a set  $\Omega$  is a set of subsets of  $\Omega$  containing  $\Omega$  and closed under taking complements and countable unions. In probability theory the elements  $F$  in  $\mathcal{F}$  represent events.

**Generated  $\sigma$ -algebra:** The smallest  $\sigma$ -algebra on a set  $\Omega$  containing a set  $\mathcal{G}$  of subsets of  $\Omega$  is called the  $\sigma$ -algebra generated by  $\mathcal{G}$  and denoted by  $\sigma(\mathcal{G})$ .

**Borel  $\sigma$ -algebra:** The  $\sigma$ -algebra generated by all open sets in  $\mathbb{R}$  is called the Borel  $\sigma$ -algebra on  $\mathbb{R}$  and denoted by  $\mathcal{B}(\mathbb{R})$ .

**Measurable space:** A pair  $(\Omega, \mathcal{F})$  where  $\mathcal{F}$  is a  $\sigma$ -algebra on  $\Omega$  is called a measurable space.

**Measure:** A measure on a measurable space  $(\Omega, \mathcal{F})$  is a function  $\mu : \mathcal{F} \rightarrow \mathbb{R}^+ \cup \{\infty\}$  such that

$$\begin{aligned}\mu(\emptyset) &= 0 \\ \mu\left(\bigcup_{i=1}^{\infty} A_i\right) &= \sum_{i=1}^{\infty} \mu(A_i)\end{aligned}$$

for a countable set  $\{A_i\}_{i=1}^{\infty}$  of disjoint members of  $\mathcal{F}$ .

**Semi-ring:** A semi-ring  $S$  on a set  $\Omega$  is a set of subsets of  $\Omega$  such that

- $\emptyset \in \mathcal{F}$
- $A \cap B \in \mathcal{F}$  for all  $A, B \in \mathcal{F}$
- for all  $A, B \in \mathcal{F}$  there exist a finite number  $K_i \in \mathcal{F}$ ,  $1 \leq i \leq n$ , such that  

$$A \setminus B = \bigcup_{i=1}^n K_i$$

**pre-measure:** A pre-measure on a semi-ring  $(\Omega, S)$  is a function  $\mu : S \rightarrow \mathbb{R}^+ \cup \{\infty\}$  such that

$$\begin{aligned}\mu(\emptyset) &= 0 \\ \mu\left(\bigcup_{i=1}^{\infty} A_i\right) &= \sum_{i=1}^{\infty} \mu(A_i)\end{aligned}$$

for each countable set  $\{A_i\}_{i=1}^{\infty} \in S$  of disjoint members of  $S$ . Moreover  $\mu$  is said to be  $\sigma$ -finite if  $\Omega$  is a countable union of sets  $A_i$  in  $S$  with  $\mu(A_i) < \infty$ .

**Caratheodory's extension theorem:** Let  $\mu$  be a  $\sigma$ -finite pre-measure on a semi-ring  $(\Omega, S)$ . Then it uniquely extends to a measure on its generated sigma algebra  $\sigma(S)$ .

**Measure space:** A measure space is a triple  $(\Omega, \mathcal{F}, \mu)$  consisting of a measurable space  $(\Omega, \mathcal{F})$  and a measure  $\mu$  on this space.

**Probability measure:** A probability measure  $P$  on a measurable space  $(\Omega, \mathcal{F})$  is a measure on  $(\Omega, \mathcal{F})$  such that  $P(\Omega) = 1$ .

**Probability space:** A probability space is a triple  $(\Omega, \mathcal{F}, P)$  consisting of a measurable space  $(\Omega, \mathcal{F})$  and a probability measure  $P$  on this space.

**Measurable function:** A measurable function  $f : (\Omega_1, \mathcal{F}_1) \rightarrow (\Omega_2, \mathcal{F}_2)$  from one measurable space to another is a function  $f : \Omega_1 \rightarrow \Omega_2$  such that

$$f^{-1}(F) \in \mathcal{F}_1 \text{ for all } F \in \mathcal{F}_2.$$

**Lebesgue integral:** The Lebesgue integral with respect to a measure  $\mu$  on a measurable space  $(\Omega, \mathcal{F})$  is a functional from a subset of measurable functions  $f : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R} \cup \{-\infty, \infty\}, \mathcal{B}(\mathbb{R} \cup \{-\infty, \infty\}))$  having a lot of nice properties. The Lebesgue integral of  $f$  is often denoted by

$$\int_{\Omega} f d\mu \text{ or } \int_{\Omega} f(\omega) d\mu(\omega).$$

Note however that the Lebesgue integral is not defined for every measurable function. In case the Lebesgue integral of  $|f|$  is defined and not equal to  $\infty$ , we say that  $f$  is Lebesgue integrable. Here we do not elaborate further on the construction of the Lebesgue integral, but just give its most important properties:

- the Lebesgue integral of any indicator function equals its measure:

$$\int_{\Omega} 1_A d\mu = \mu(A)$$

for all  $A \in \mathcal{F}$ .

- linearity:

$$a \int_{\Omega} f d\mu + b \int_{\Omega} g d\mu = \int_{\Omega} (af + bg) d\mu$$

for  $a, b \in \mathbb{R}$  and  $f, g$  Lebesgue integrable functions.

- monotonicity:

$$\int_{\Omega} f d\mu \leq \int_{\Omega} g d\mu$$

for non-negative measurable functions

$$f \leq g$$

- monotone convergence: Let  $\{f_k\}$  be a series of increasing non-negative measurable functions, then

$$\lim \int_{\Omega} f_k d\mu = \int_{\Omega} \lim f_k d\mu$$

- Fatou's lemma: Let  $\{f_k\}$  be a series of non-negative measurable functions, then

$$\int_{\Omega} \liminf_k f_k d\mu \leq \liminf_k \int_{\Omega} f_k d\mu$$

- dominated convergence theorem: Let  $\{f_k\}$  be a series of measurable functions with pointwise limit  $f$  such that for all  $k$ ,  $|f_k| \leq g$  for a Lebesgue integrable function  $g$ , then  $f$  is Lebesgue integrable and

$$\lim_k \int_{\Omega} f_k d\mu = \int_{\Omega} f d\mu$$

**Random variable:** A random variable  $X$  on a probability space  $(\Omega, \mathcal{F}, P)$  is a measurable function  $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

**Almost sure:** In a probability space  $(\Omega, \mathcal{F}, P)$  an event  $F \in \mathcal{F}$  is said to happen almost surely if  $P(F) = 1$ .

**$\sigma$ -algebra generated by a random variable:** The  $\sigma$ -algebra  $\{X^{-1}(U) | U \in \mathcal{B}(\mathbb{R})\}$  is called the  $\sigma$ -algebra generated by  $X$  and denoted by  $\sigma(X)$ .

**Independence:** Two random variables  $X, Y$  on a probability space  $(\Omega, \mathcal{F}, P)$  are independent if

$$P(A \cap B) = P(A)P(B)$$

for every  $A \in \sigma(X), B \in \sigma(Y)$ .

**Expectation:** Let  $X$  be a random variable on a probability space  $(\Omega, \mathcal{F}, P)$  and  $g : \mathbb{R} \rightarrow \mathbb{R}$ . We define the expectation of  $g(X)$

$$E(g(X)) = \int_{\Omega} g(X)dP,$$

if this Lebesgue integral exists.

**Conditional expectation:** Let  $X$  be a random variable on a probability space  $(\Omega, \mathcal{F}, P)$  and let  $\mathcal{H}$  be a sub  $\sigma$ -algebra of  $\mathcal{F}$ . A conditional expectation  $E(X|\mathcal{H})$  of  $X$  with respect to  $\mathcal{H}$  is a random variable on  $(\Omega, \mathcal{H})$  defined through the following property

$$\int_H E(X|\mathcal{H})dP = \int_H XdP$$

for every  $H \in \mathcal{H}$ . It is unique up to a set of measure 0. Let  $Y$  be another random variable on this probability space. The conditional expectation  $E(X|Y)$  of  $X$  conditional on  $Y$  is defined by

$$E(X|Y) = E(X|\sigma(Y)).$$

**Almost sure convergence:** A sequence of random variables  $(X_n)_{n=1}^{\infty}$  on a probability space  $(\Omega, \mathcal{F}, P)$  converges almost surely to a random variable  $X$  on the same space if

$$P\left(\lim_{n \rightarrow \infty} X_n \neq X\right) = 0.$$

**Convergence in probability:** A sequence of random variables  $(X_n)_{n=1}^{\infty}$  on a probability space  $(\Omega, \mathcal{F}, P)$  converges in probability to a random variable  $X$  on the same space if

$$\forall \epsilon > 0 : \lim_{n \rightarrow \infty} P\{|X_n - X| > \epsilon\} = 0.$$

**Convergence in mean square:** We say that a sequence of random variables  $(X_n)_{n=1}^{\infty}$  on a probability space  $(\Omega, \mathcal{F}, P)$  converges in mean square to a random variable  $X$  on the same space if

$$\lim_{n \rightarrow \infty} E((X_n - X)^2) = 0.$$

**Exercise 1.** Give the smallest  $\sigma$ -algebra  $\mathcal{F}$  on the set  $\Omega = \{1, 2, 3, 4, 5, 6\}$ .

**Exercise 2.** Give the largest  $\sigma$ -algebra  $\mathcal{F}$  on the set  $\Omega = \{1, 2, 3, 4, 5, 6\}$ . It can be used to describe all events of throwing a regular dice.

**Exercise 3.** Give the  $\sigma$ -algebra  $\mathcal{F}$  on the set  $\Omega = \{1, 2, 3, 4, 5, 6\}$  generated by  $\{\{1, 3, 5\}\}$ . What kind of events in throwing a regular dice can it describe?

**Exercise 4.** Show that a  $\sigma$ -algebra is closed under countable intersections

**Exercise 5.** Give the smallest  $\sigma$ -algebra  $\mathcal{F}$  on the set  $\mathbb{R}$ .

**Exercise 6.** Give the smallest  $\sigma$ -algebra  $\mathcal{F}$  on the set  $\mathbb{R}$  that can be used to describe whether a real number is (strictly) positive or not.

**Exercise 7.** How does the smallest  $\sigma$ -algebra  $\mathcal{F}$  on the set  $\mathbb{R}$  that can be used to check whether a real number lies in an open interval  $]a, b[$  for all real numbers  $a \leq b$  look like? Show it must contain all intervals  $]a, b[, [a, \infty[, [-\infty, a], [a, b], ]a, b], [a, b[, \mathbb{N}, \mathbb{Z}, \mathbb{Q}$ . Show that  $\mathcal{F}$  is the Borel sigma algebra on  $\mathbb{R}$ .

**Exercise 8.** Describe a probability space  $(\Omega, \mathcal{F}, P)$  representing throws of a fair regular dice.

**Exercise 9.** Give an example of a probability space  $(\mathbb{N}, \mathcal{P}(\mathbb{N}), P)$  such that  $P(\{n\}) > 0$  for all  $n \in \mathbb{N}$ . Note that it suffices to define  $P$  on all singletons, since every subset  $S$  of  $\mathbb{N}$  is a countable union of singletons. So its probability measure is implicitly defined: it equals the sum of the measures of all singletons in  $S$  and this sum exists as it is a sum of positive integers bounded by  $P(\mathbb{N}) = 1$ . Also note that it is impossible to take  $P(\{n\})$  equal for all  $n$ . So drawing a random uniform natural number is impossible. In order to draw random natural numbers you hence need to specify a probability space.

**Exercise 10.** It is challenging to find a measure space  $(\mathbb{R}, \mathcal{F}, \mu)$  with the following natural properties:  $\mathcal{F}$  contains all intervals,  $\mu([0, 1]) = 1$ ,  $\mu(A) = \mu(A + x)$  for all  $x \in \mathbb{R}$  and  $A \in \mathcal{F}$  (translation invariant),  $A \in \mathcal{F}$  for all  $A \subset B$  for all  $B \in \mathcal{F}$  with  $\mu(B) = 0$  (completeness). You may try to define such measure space but don't get frustrated: it is a deep result in measure theory. The Lebesgue measure space  $(\mathbb{R}, \mathcal{F}, \mu)$  is the unique space satisfying these properties. Its sigma algebra contains all intervals  $(a, b)$  for  $b > a$  and their measure equals  $\mu((a, b)) = b - a$ . It is not evident to construct it and prove it satisfies the properties of  $\sigma$ -algebra and measure. It does contain the Borel  $\sigma$ -algebra  $\mathcal{B}(\mathbb{R})$ . In fact it is a bit larger than  $\mathcal{B}(\mathbb{R})$ . The Borel  $\sigma$ -algebra  $\mathcal{B}(\mathbb{R})$  contains sets of Lebesgue measure 0 that have subsets not in  $\mathcal{B}(\mathbb{R})$ . The Lebesgue  $\sigma$ -algebra is then the completion of the Borel  $\sigma$ -algebra in the sense that it is the smallest  $\sigma$ -algebra containing the Borel  $\sigma$ -algebra and all subsets of borel sets with measure 0. For more information see Lebesgue measure.

**Exercise 11.** A famous paradox in probability theory is Bertrand's paradox: Consider an equilateral triangle inscribed in a circle. When a chord of the circle is chosen at random, what's the probability that the chord is longer than a side of the triangle? You can compute the probability in three different ways yielding three different results. That's the paradox. See Bertrand paradox The problem is that it is not well defined what is meant with choosing a random chord of the circle. In each of the three solutions a different probability space is considered: describe this space in each of the cases.

**Exercise 12.** Give an explicit description of a standard normal random variable  $X$  and its probability space  $(\Omega, \mathcal{F}, P)$

**Exercise 13.** Let  $(\Omega, \mathcal{F}, P)$  be a probability space on which two sub  $\sigma$ -algebras  $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$  are defined. For a random variable  $X$  on such a space, prove the tower law that states that if  $E[X]$  is defined, then

$$E[E[X|\mathcal{G}_2]|\mathcal{G}_1] = E[X|\mathcal{G}_1] \quad (\text{a.s.})$$

We often use the tower law for a stochastic process with a filtration  $(\mathcal{F}_t)$  in the sense that  $E[g(X_T)] = E[E[g(X_T)|\mathcal{F}_t]]$

In the definition of conditional expectation we condition on a random variable (i.e. the  $\sigma$ -algebra it generates)  $E[X|Y]$ . In particular we do not define such thing like  $E[X|Y = y]$ . This is because the event  $Y = y$  might have 0 probability and in such cases we cannot define it in a consistent manner as explained by the Borel-Kolmogorov paradox

**Exercise 14.** Show that almost sure convergence implies convergence in probability and convergence in mean square implies convergence in probability. See also Convergence of random variables.

### 2.1.2 Stochastic Processes

**Stochastic process:** A stochastic process on a probability space  $(\Omega, \mathcal{F}, P)$  is a collection of random variables  $(X_t)_{t \in \mathbb{R}^+}$  on that space. A stochastic process can also be seen as a function  $X : \mathbb{R} \times \Omega \rightarrow \mathbb{R}$  where each  $X_t : \Omega \rightarrow \mathbb{R} : \omega \mapsto X(t, \omega)$  is a random variable on the probability space  $(\Omega, \mathcal{F}, P)$ . The functions  $X(-, \omega) : \mathbb{R}^+ \rightarrow \mathbb{R} : t \mapsto X(\omega, t)$  for  $\omega \in \Omega$  are called the sample paths of the process.

**Filtered probability space:** A filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}^+})$  is a probability space  $(\Omega, \mathcal{F}, P)$  together with a filtration  $(\mathcal{F}_t)_{t \in \mathbb{R}^+}$ , i.e. an indexed set of  $\sigma$ -algebras such that

$$\mathcal{F}_t \subset \mathcal{F}_s \text{ for all } t \leq s.$$

The filtration represents the information available up to time  $t$ .

**Adapted process:** Given a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}^+})$  a stochastic process  $X_t$  on  $(\Omega, \mathcal{F}, P)$  is adapted to the filtration if

$$\forall t \in \mathbb{R}^+, \forall A \in \mathcal{B}(\mathbb{R}) : X_t^{-1}(A) \in \mathcal{F}_t.$$

Intuitively an adapted process is one that cannot see into the future.

**Predictable process:** A stochastic process  $X$  on a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}^+})$  is a predictable process if it is a measurable function from  $(\mathbb{R} \times \Omega, \Sigma)$  to  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , where  $\Sigma$  denotes the  $\sigma$ -algebra generated by all left continuous adapted processes on the filtered probability space. Intuitively a predictable process is one that is predictable from prior times. Note that adapted left continuous processes clearly satisfy the definition and are predictable in this intuitive sense as its value at a given time  $t$  is (a.s.) equal to the limit of its values at prior times.

**Partition:** a sequence  $x = (x_0, \dots, x_n)$  of real numbers with  $a = x_0 < x_1 < \dots < x_n = b$  is called a partition of the interval  $[a, b]$ .

**Total variation:** The total variation  $V_a^b(f)$  of a function  $f : [a, b] \rightarrow \mathbb{R}$  is defined as

$$\sup_{x \in \mathcal{X}} \sum_{i=1}^{N_x} |f(x_i) - f(x_{i-1})|,$$

where

$$\mathcal{X} = \{x = (x_1, \dots, x_{N_x}) | x \text{ is a partition of } [a, b]\}.$$

**Finite variation:** A function  $f : \mathbb{R}^+ \rightarrow \mathbb{R}$  has bounded (or finite) variation if its total variation over every finite interval is finite. Otherwise it has unbounded variation. A random process  $X$  has finite variation if (almost) all its sample paths have finite variation.

**Norm (or mesh) of a partition:** The norm  $|x|$  of a partition  $x = (x_0, x_1, \dots, x_n)$  is the maximum of its increments

$$|x| = \max_{i=1}^n (x_i - x_{i-1}).$$

**Quadratic variation:** The quadratic variation  $[X]_t$  of a stochastic process  $X$  on a probability space  $(\Omega, \mathcal{F}, P)$  is defined as

$$[X]_t = \lim_{|\pi| \rightarrow 0} \sum_{i=1}^n (X(t_i) - X(t_{i-1}))^2,$$

where  $\pi = (t_0, t_1, \dots, t_n)$  runs over partitions of  $[0, t]$  and the limit is taken with respect to convergence in probability.

**Quadratic covariation:** The quadratic covariation  $[X, Y]_t$  of 2 stochastic processes  $X$  and  $Y$  on a probability space  $(\Omega, \mathcal{F}, P)$  is defined as

$$[X, Y]_t = \lim_{|\pi| \rightarrow 0} \sum_{i=1}^n (X(t_i) - X(t_{i-1}))(Y(t_i) - Y(t_{i-1})),$$

where  $\pi = (t_0, t_1, \dots, t_n)$  runs over partitions of  $[0, t]$  and the limit is taken with respect to convergence in probability.

**Càdlàg and càglàd:** a càdlàg, resp. càglàd, stochastic process  $X$  is a process whose sample paths are right continuous with left limits (continue à droite, limite à gauche), resp. left continuous with right limits (continue à gauche, limite à droite).

**Martingale:** Given a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}^+})$  an adapted stochastic process  $X_t$  is a martingale if its sample paths are almost surely càdlàg and for each  $s$  and  $t > s$  we have

$$E(|X(t)|) < \infty$$

and

$$E(X(t)|\mathcal{F}_s) = X(s).$$

**Stopping Time:** Given a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}^+})$ , a random variable  $\tau : \Omega \rightarrow \mathbb{R}^+$  is a stopping time if for all  $t \in \mathbb{R}^+$

$$\{\omega \in \Omega | \tau(\omega) < t\} \in \mathcal{F}_t.$$

**Local Martingale:** Given a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}^+})$ . A stochastic process  $X_t$  adapted to this filtration is a local martingale if there exists a sequence of stopping times  $\tau_k$  such that

- the  $\tau_k$  are almost surely increasing:  $P(\tau_k < \tau_{k+1}) = 1$
- the  $\tau_k$  diverge almost surely:  $P(\lim_{k \rightarrow \infty} \tau_k = \infty) = 1$
- the stopped processes  $X_{\min(\tau_k, t)}$  are martingales for all  $k$

**Locally bounded:** Given a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}^+})$ . A stochastic process  $X_t$  is a locally bounded process if there exists a sequence of stopping times  $\tau_k$  such that

- the  $\tau_k$  are almost surely increasing:  $P(\tau_k < \tau_{k+1}) = 1$
- the  $\tau_k$  diverge almost surely:  $P(\lim_{k \rightarrow \infty} \tau_k = \infty) = 1$
- the stopped processes  $X_{\min(\tau_k, t)}$  is bounded for all  $k$

**Semimartingale:** Given a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t)_{t \in \mathbb{R}^+})$ . A semi-martingale is a stochastic process on this space that is the sum of a martingale and a càdlàg adapted process with finite variation.

**Exercise 15.** Find a function with unbounded variation.

**Exercise 16.** Find a continuous function with unbounded variation but bounded quadratic variation.

**Exercise 17.** Prove that a continuously differentiable function has 0 quadratic variation.

## 2.2 Brownian Motion

A Brownian motion on a probability space  $(\Omega, \mathcal{F}, P)$  is defined through its properties. It is a stochastic process  $B(t)$  on the space such that

- $B(0) = 0$
- $B(t) - B(s)$  has a normal  $N(0, t - s)$  distribution for  $s \leq t$
- its increments over disjoint time intervals are independent
- its paths are almost surely continuous

Note that it requires some work to prove that Brownian motion exists. Moreover one can show the following properties

- Brownian motion is almost surely nowhere differentiable
- Brownian motion has almost surely unbounded variation
- Brownian motion has quadratic variation  $[B]_t = t$ .

The construction of Brownian motion is based on the following property of normal distributions. Given

- $X_1 \sim N(0, \sigma_1^2)$
- $X_2 \sim N(0, \sigma_2^2)$
- $X_1, X_2$  are independent

then

$$X = X_1 + X_2 = N(0, \sigma_1^2 + \sigma_2^2).$$

We can use this idea to construct Brownian motion on a discrete time grid  $[0, \Delta, 2\Delta, \dots, N\Delta]$ . Choose  $N$  independent random variables  $X_i \sim N(0, \Delta)$ . Set

$$\begin{aligned} B(0) &= 0 \\ B(i\Delta) &= B((i-1)\Delta) + X_i. \end{aligned}$$

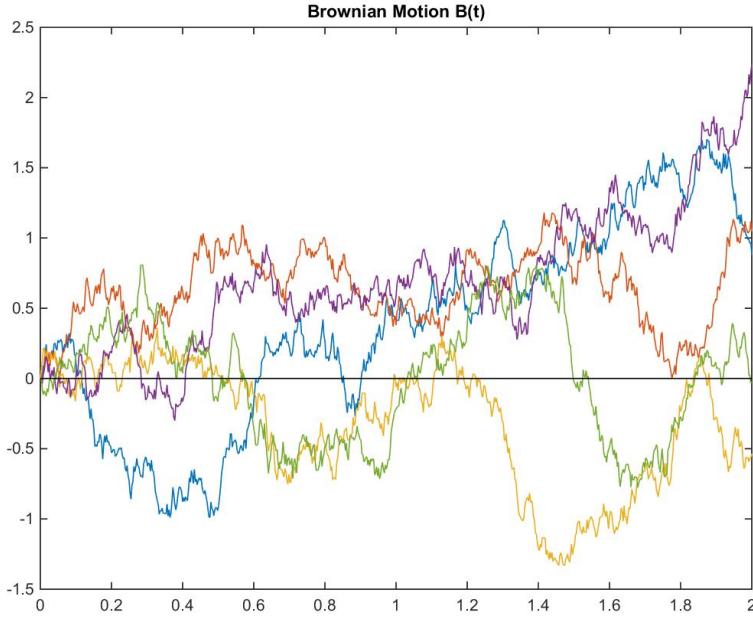


Figure 2.1: 5 sample paths of Brownian motion

### 2.3 Itô integral with respect to Brownian motion

Many processes in finance are defined by integrating Brownian motion. More precisely we need integrals like

$$\int_0^t f(s, B(s))ds$$

and

$$\int_0^t f(s, B(s))dB(s)$$

for a smooth function  $f$ . The first integral can be defined as a pathwise Riemann-Stieltjes integral since Brownian motion is (a.s.) continuous. In general, the second integral cannot be defined as a pathwise integral since the integrator  $B(s)$  is not of bounded variation. However it is defined as the limit in probability of sums of random variables over a partition

$$\int_0^t f(s, B(s))dB(s) = \lim_{|\Pi| \rightarrow 0} Q_\Pi(t)$$

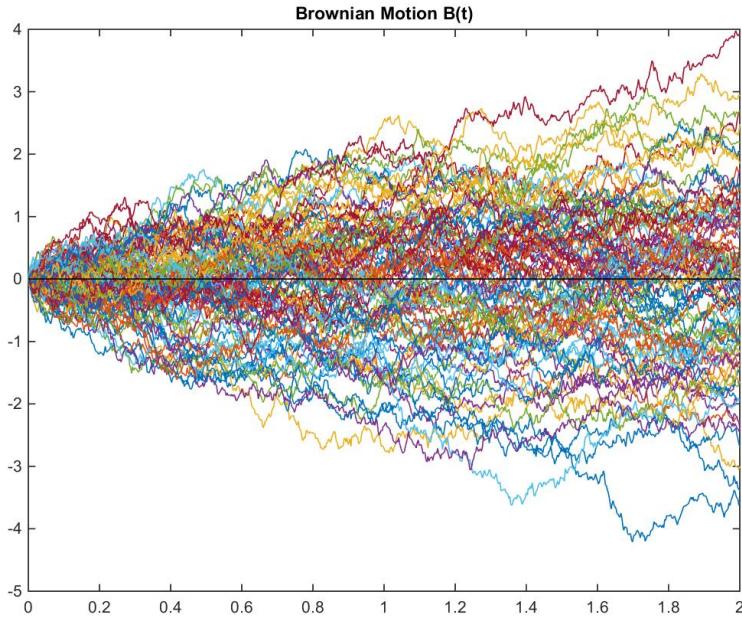


Figure 2.2: 100 sample paths of Brownian motion

where

$$Q_{\Pi}(t) = \sum_{i=1}^N f(t_{i-1}, B(t_{i-1}))(B(t_i) - B(t_{i-1})),$$

$\Pi = \{t_0 = 0, t_1, \dots, t_N = t\}$  is a partition of  $[0, t]$  and  $|\Pi| = \max_i |t_i - t_{i-1}|$  is a norm on the partition. It is a result of stochastic calculus theory that this limit exists and we call it the Itô integral. Note that whenever  $f$  satisfies

$$\int_0^t E(f(s, B(s))^2) ds < \infty$$

the limit can be taken in the mean square sense (see [Mik98]). Ito first defined the integral as a pathwise integral requiring the same condition and fast enough convergence of the norm of the partition (see [Ito44]). De La Vega showed that the limit of the partial sums does not need to converge pathwise in case of slower convergence speed (see [Veg74]).

The fact that we evaluate  $f$  in the left hand point of the partition intervals is crucial for the Itô integral and yields a martingale (as the increment of Brownian motion over a time interval does not depend on its value at the start of the interval).

As an example consider

$$\int_0^t B(s) dB(s).$$

Taking a uniformly spaced partition  $\Pi = \{t_0 = 0, t_1 = t/N, t_2 = 2t/N, \dots, t_N = t\}$  and denoting  $\Delta B_i = B(t_i) - B(t_{i-1})$  we have

$$\begin{aligned} S_N(t) &:= \sum_{i=1}^N B(t_{i-1})(B(t_i) - B(t_{i-1})) \\ &= \sum_{i=1}^N \sum_{k=1}^{i-1} \Delta B_k \Delta B_i \\ &= \frac{1}{2} \sum_{i=1}^N \sum_{k < i} \Delta B_k \Delta B_i \\ &= \frac{1}{2} \left( \sum_{i=1}^N \sum_{k=1}^N \Delta B_k \Delta B_i - \sum_{i=1}^N (\Delta B_i)^2 \right) \\ &= \frac{1}{2} \left( (\sum_{i=1}^N \Delta B_i)^2 - \sum_{i=1}^N (\Delta B_i)^2 \right) \\ &= \frac{1}{2} \left( B(t)^2 - \frac{t}{N} \sum_{i=1}^N \frac{N(\Delta B_i)^2}{t} \right) \end{aligned}$$

Now the sum in the right hand term is a sum of independent squares of standard normal random variables. Due to the law of large numbers its average converges in probability to the expectation of one such square (that equals 1). Hence

$$\int_0^t B(s) dB(s) = \frac{1}{2} (B(t)^2 - t)$$

or

$$B(t)^2 = 2 \int_0^t B(s) dB(s) + t.$$

Our computation also shows that the quadratic variation of Brownian motion, given it exists, equals  $[B(t)] = t$  (in many text books this is loosely written as  $(dB)^2 = dt$ ). The resulting equation is often also written as

$$d(B(t)^2) = 2B(t)dB(t) + d[B(t)].$$

Note that for deterministic differentiable functions  $f(t)$  (with  $f(0) = 0$ ) the fundamental theorem of calculus says

$$f(t)^2 = 2 \int_0^t f(s) df(s).$$

Hence the second term on the right-hand side is typical for stochastic calculus. We can easily check it must be there by taking expectations. Indeed,

$$\begin{aligned} E(B(t)^2) &= 2E\left(\int_0^t B(s)dB(s)\right) + t \\ &= 2 \int_0^t E(B(s))E(dB(s)) + t \\ &= t \end{aligned}$$

Here we used the fact that  $B(s)$  and  $dB(s)$  are independent.

This equation must indeed hold since  $B(t) \sim N(0, t)$ . Thus the second term is indeed necessary.

**Exercise 18.** Show that the quadratic variation of Brownian motion exists and equals  $[B(t)] = t$ . You need to consider arbitrary partitions, instead of the equally spaced partitions we worked with above. Hint: show that the partition sums converge in mean square sense.

**Exercise 19.** Implement code to generate sample paths of Brownian motion at a discrete time grid

**Exercise 20.** Show that Brownian motion has almost surely unbounded variation. Hint: This is quite hard to prove directly and therefore often proved by contradiction. Hence assume Brownian motion has bounded variation on some interval  $[a, b]$  (with  $b > a$ ) and a set  $A \subset \Omega$  with  $P(A) > 0$ . Then show it must have 0 quadratic variation on  $[a, b]$  on a subset  $B \subset \Omega$  with  $P(B) > 0$ , in contradiction with its known quadratic variation  $b - a > 0$ .

**Exercise 21.** Let  $\hat{B}(t) := \sup_{0 \leq s \leq t} B(s)$  for a Brownian motion  $B(t)$ . Prove that  $P(\hat{B}(t) \geq b) = 2P(B(t) \geq b)$  for  $b \geq 0$ . This is a nice property of Brownian motion that can be exploited in pricing barrier options. It uses the reflection principle, see for instance maximum of Brownian motion

**Exercise 22.** Suppose that we replace  $Q_\Pi(t)$  by

$$Q_\Pi^a(t) = \sum_{i=1}^N ((1-a)f(t_{i-1}, B(t_{i-1})) + af(t_i, B(t_i))) (B(t_i) - B(t_{i-1}))$$

in the definition of the stochastic integral. Show that then

$$\int_0^t B(s)dB(s) = \frac{1}{2} (B(t)^2 - t) + at.$$

What happens when  $a = \frac{1}{2}$ ?

**Exercise 23.** Two-dimensional Brownian motion is defined similarly as 1-dimensional Brownian motion. it is a 2-dimensional stochastic process  $(B_1(t), B_2(t))$  on a probability space such that

- $B_1(0) = B_2(0) = 0$
- $(B_1(t) - B_1(s), B_2(t) - B_2(s))$  has a 2-dimensional normal distribution  $N(0, (t-s)C)$  for  $s \leq t$  where  $C$  denotes the correlation matrix

$$C = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

- its increments over disjoint time intervals are independent
- its paths are almost surely continuous

Show that the quadratic covariation  $[B_1, B_2](t) = \rho t$  (in many text books this is written as  $dB_1 dB_2 = \rho dt$ ) and show that

$$d(B_1(t)B_2(t)) = B_1(t)dB_2(t) + B_2(t)dB_1(t) + \rho dt.$$

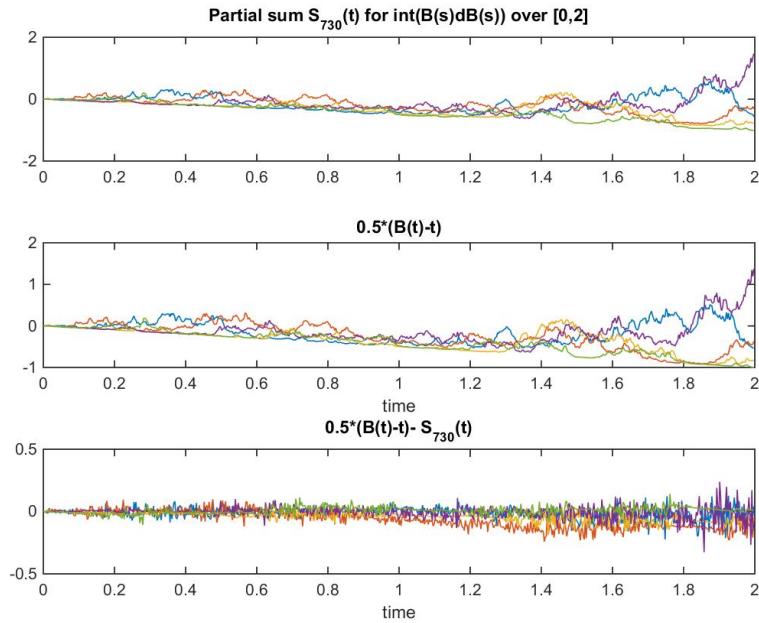


Figure 2.3: Integrating Brownian motion

**Itô Isometry:** The Itô integral satisfies the isometry property

$$\mathbb{E} \left[ \left( \int_0^t f(s, B(s)) dB(s) \right)^2 \right] = \mathbb{E} \left[ \int_0^t f(s, B(s))^2 ds \right].$$

In particular, for a deterministic function  $f(s)$  we have

$$\mathbb{E} \left[ \left( \int_0^t f(s) dB(s) \right)^2 \right] = \int_0^t f(s)^2 ds.$$

It follows that

$$\int_0^t f(s) dB(s) \sim N(0, \int_0^t f(s)^2 ds).$$

## 2.4 Itô's Lemma

**Lemma:** Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a continuous function, then

$$\sum_{i=1}^N f(B(t_{i-1}))(B(t_i) - B(t_{i-1}))^2$$

converges in probability to

$$\int_0^T f(B(t)) dt$$

when  $N \rightarrow \infty$ . Here  $t_i = iT/N$ .

Proof: First assume  $|f|$  is bounded by some constant  $C$ . It suffices to prove that

$$I_n = \sum_{i=1}^N f(B(t_{i-1})) ((B(t_i) - B(t_{i-1}))^2 - (t_i - t_{i-1}))$$

converges in probability to 0. To do so we show that  $\mathbb{E}[I_n^2]$  converges to 0. Denote  $f(B(t_i))$  by  $f_i$ ,  $B(t_i) - B(t_{i-1})$  by  $\Delta B_i$ ,  $t_i - t_{i-1}$  by  $\Delta t_i$ . Then

$$\begin{aligned} \mathbb{E}[I_n^2] &= \sum_{i=1}^N \sum_{j=1}^N \mathbb{E}[f_{i-1} f_{j-1} ((\Delta B_i)^2 - \Delta t_i)((\Delta B_j)^2 - \Delta t_j)] \\ &= \sum_{i=1}^N \mathbb{E}[f_{i-1}^2] \mathbb{E}[(\Delta B_i)^2 - \Delta t_i]^2 \\ &\leq C^2 \sum_{i=1}^N (\Delta t_i)^2 \mathbb{E}[(B_1^2 - 1)^2] \\ &\leq C^2 \mathbb{E}[(B_1^2 - 1)^2] |\Pi| T \end{aligned}$$

Note here that the expectation over the cross terms equals zero since increments of Brownian motion over a time interval are independent from the value at the start of the interval. This expectation clearly converges to 0 as  $N \rightarrow \infty$ . In the general case when  $f$  is not necessarily bounded, we may replace  $f$  by  $f\psi$  where  $\psi(x)$  is a smooth function that equals 1 if  $|x| \leq M$  and 0 if  $|x| \geq M + 1$

for some constant  $M$ . The lemma then holds on the subset of paths for which  $|B(t)| \leq M$  on  $[0, T]$ . As  $M \rightarrow \infty$  the probability of this subset goes to 1. Hence in the unbounded case convergence in probability holds.  $\square$

**Exercise 24.** Show that the sums

$$I_n = \sum_{i=1}^N f(B(t_{i-1})) ((B(t_i) - B(t_{i-1}))^2 - (t_i - t_{i-1}))$$

converge in mean square sense when

$$\int_0^T E(f(t)^2) dt < \infty.$$

**Itô's Lemma:** Let  $f : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}$  be a function in two variables  $x$  and  $t$  with continuous first and second order partial derivatives. Let  $B_t$  be a Brownian motion on a probability space  $(\Omega, \mathcal{F}, P)$ . Then for all  $T > 0$

$$f(B(T), T) - f(0, 0) = \int_0^T f_t(B(t), t) dt + \int_0^T f_x(B(t), t) dB(t) + \frac{1}{2} \int_0^T f_{xx}(B(t), t) dt.$$

Proof:

We first prove the lemma in case  $f$  also has a continuous third order partial derivatives. Using the same partition and notation as before ( $f_{x,i} = f_x(B(t_i), t_i)$ , ...) and using Taylor's expansion on the set  $\Omega_M = \{\omega \in \Omega | \forall t \in [0, T] : |B_t(\omega)| \leq M\}$  we have

$$f(B(T), T) - f(0, 0) = \sum_{i=1}^N f_i - f_{i-1} \quad (2.1)$$

$$= \sum_{i=1}^N f_{t,i} \Delta t_i \quad (2.2)$$

$$+ \sum_{i=1}^N f_{x,i} \Delta B_i \quad (2.3)$$

$$+ \frac{1}{2} \sum_{i=1}^N f_{xx,i} \Delta B_i^2 \quad (2.4)$$

$$+ \frac{1}{2} \sum_{i=1}^N R_{tx}(t_i, B(t_i)) \Delta t_i \Delta B_i \quad (2.5)$$

$$+ \frac{1}{2} \sum_{i=1}^N R_{tt}(t_i, B(t_i)) (\Delta t_i)^2 \quad (2.6)$$

$$+ \frac{1}{6} \sum_{i=1}^N R_{xxx}(t_i, B(t_i)) (\Delta B_i)^3 \quad (2.7)$$

$$+ \frac{1}{6} \sum_{i=1}^N R_{txx}(t_i, B(t_i)) \Delta t_i (\Delta B_i)^2 \quad (2.8)$$

$$+ \frac{1}{6} \sum_{i=1}^N R_{ttx}(t_i, B(t_i)) (\Delta t_i)^2 \Delta B_i \quad (2.9)$$

$$+ \frac{1}{6} \sum_{i=1}^N R_{ttt}(t_i, B(t_i)) (\Delta B_i)^3 \quad (2.10)$$

where the remainder coefficients  $R_\beta(t_{i-1}, B(t_{i-1}))$  satisfy

$$\begin{aligned} |R_\beta(t_i, B(t_i))| &\leq \max_{|\alpha|=3} \max_{t_{i-1} \leq t \leq t_i} \max_{-2M \leq x \leq 2M} |f_\alpha(t, x)| \\ &\leq \max_{|\alpha|=3} \max_{0 \leq t \leq T} \max_{-2M \leq x \leq 2M} |f_\alpha(t, x)| \\ &= K_{|\beta|} < \infty \end{aligned}$$

Here  $|\alpha|$  denotes the order of the multi-index and the  $K_i < \infty$  for  $i \leq 3$  because up to the third order the partial derivatives are assumed continuous and hence they are bounded on the compact set  $[0, T] \times [-2M, 2M]$

We then have

$$\begin{aligned}
\left| \sum_{i=1}^N R_{tx}(t_i, B(t_i)) \Delta t_i \Delta B_i \right| &\leq K_2 T \sup_i |\Delta B_i| \\
\left| \sum_{i=1}^N R_{tt}(t_i, B(t_i)) (\Delta t_i)^2 \right| &\leq K_2 T \sup_i |\Delta t_i| \\
\left| \sum_{i=1}^N R_{xxx}(t_i, B(t_i)) (\Delta B_i)^3 \right| &\leq K_3 \sup_i |\Delta B_i| \sum_{i=1}^N \Delta B_i^2 \\
\left| \sum_{i=1}^N R_{txx}(t_i, B(t_i)) \Delta t_i (\Delta B_i)^2 \right| &\leq K_3 \sup_i |\Delta t_i| \sum_{i=1}^N \Delta B_i^2 \\
\left| \sum_{i=1}^N R_{txx}(t_i, B(t_i)) (\Delta t_i)^2 \Delta B_i \right| &\leq K_3 T \sup_i |\Delta t_i| \sup_i |\Delta B_i| \\
\left| \sum_{i=1}^N R_{ttt}(t_i, B(t_i)) (\Delta t_i)^3 \right| &\leq K_3 T (\sup_i |\Delta t_i|)^2
\end{aligned}$$

The terms 2.5, ... 2.10 then converge to 0 in probability due to the continuity of  $B(t)$ . The first term 2.2 converges to the pathwise Riemann integral and the second term 2.3 to the Itô integral. For the third term 2.4 we may use our previous lemma yielding the required result.

This proves the lemma for functions with third order continuous derivatives on  $\Omega_M$ . However, the set  $\Omega_M$  converges to a set of  $P$ -measure 1. Hence the lemma holds a.s. for such functions.

We now turn to the general case where we only require  $f$  to be a  $C^2$  function. Due to the Stone-Weierstrass theorem we can find a series of polynomials  $g_m$  that converge, as well as their first and second order partial derivatives, to  $f$  and its first and second order partial derivatives, uniformly on the compact set  $B_n = \{(x, t) : |x| \leq n, 0 \leq t \leq T\}$ . It then follows from the properties of the Itô integral that the integrals in the Itô formula for  $g_m$  converge in probability to the integrals in the Itô formula for  $f$  on this set  $B_n$ . Since  $P(B_n) \rightarrow 1$  as  $N \rightarrow \infty$ , it follows that they also converge in probability on the whole domain. The result now follows from the first part of the proof (since polynomials have continuous third order partial derivatives).  $\square$

Note that Itô's lemma is often written in differential form (but strictly spoken this is non-sensical as  $B(t)$  is nowhere differentiable)

$$df(B(t), t) = f_t(B_t, t)dt + f_x(B_t, t)dB(t) + \frac{1}{2}f_{xx}(B(t), t)dt.$$

## 2.5 Itô processes

Here we generalise the Itô integral to integrals with respect to slightly more general processes. We say an Itô process on a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$

is any adapted stochastic process satisfying:

$$X(t) = X(0) + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dB(s), \quad (2.11)$$

where  $B(t)$  is a Brownian motion on the space and  $\mu(t)$  and  $\sigma(t)$  are predictable processes such that  $\mu(t)$  is (Lebesgue) integrable and  $\sigma(t)$  is  $B$ -integrable. Note that the definition allows  $\mu$  and  $\sigma$  to be stochastic processes. They may even depend on  $X(t)$  itself, such that the definition is implicit. We can also write it as a SDE:

$$dX(t) = \mu(t)dt + \sigma(t)dB(t).$$

Itô's lemma can then be extended to Itô processes.

**Itô's lemma for Itô processes:** Let  $f : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{R}$  be a function in two variables  $x$  and  $t$  with continuous first and second order partial derivatives. Let  $X$  be an Itô process on a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$  defined as in equation 2.11.

Then  $f(X(t), t)$  is again an Itô process and for all  $T > 0$

$$f(X(T), T) - f(0, 0) = \int_0^T f_t(X(t), t)dt + \int_0^T f_x(X(t), t)dX(t) + \frac{1}{2} \int_0^T f_{xx}(X(t), t)\sigma^2(t)dt.$$

The proof is essentially the same as before noting that for increments  $(dX(t))^2 = \sigma^2(t)(dB(t))^2 + O(dt dB(t))$ . See [Ito51] for a rigorous proof.

**Exercise 25.** Assume  $x(t)$  is a stochastic process that satisfies

$$dx(t) = -\kappa x(t)dt + \sigma dB(t).$$

Show that

$$x(t) = x(0)e^{-\kappa t} + \sigma \int_0^t e^{-\kappa(t-u)}dB(u).$$

(Hint: use Itô's lemma with  $f(x, t) = xe^{\kappa t}$ .) Next use Itô's isometry to show that

$$x(t) \sim N(x(0)e^{-\kappa t}, \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa t})).$$

## 2.6 Semimartingales

The whole theory of Ito stochastic integration extends to a very wide class of processes. This theory is developed in e.g. [JS03] where one defines

$$\int_0^t H(s)dX(s)$$

for a semimartingale  $X$  and a locally bounded predictable process  $H$ . The Ito integral extends to functions of semimartingales in higher dimensions. Note that semimartingales are not restricted to continuous functions.

# Chapter 3

# Derivative Valuation

An excellent introduction to the theory of derivative pricing is [BR96]. The fundamental asset pricing theorems were first stated by Harrison and Pliska [HP83].

## 3.1 Fundamental Asset Pricing Theorems

The use of stochastic processes in valuating instruments is based on the so-called fundamental asset pricing theorem. It states that we can compute the fair value of a derivative on an asset by computing its discounted expected payoff in the so-called risk neutral world, meaning that we need to change the probabilities of the stochastic payoff in a specific way. In such a risk neutral world, also the current fair value of the asset itself must equal the expectation of its discounted value at any future time. Often the model for the asset is specified in the risk neutral world and calibrated such that this holds. In the risk neutral world, stock prices will have an expected growth rate equal to the risk free spot rate.

Note that the payoff of more complex derivatives (such as lookback options) may depend on the value of the asset at different future times. Hence we need a stochastic process to model the asset, not just a random variable.

**Asset:** an asset is a positive adapted stochastic process  $A(t)$  on a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F})_t)$ . The process describes the stochastic value of a non-dividend paying tradable asset. Examples are the bank account or a zero coupon bond.

**Market model:** a market model  $(N, A_i)$  on a filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F})_t)$  consists of a number of tradable assets. The first asset is called the numeraire  $N(t)$  and has a special meaning: it will be used to compare the values of the other assets. The measure  $P$  is called the real world measure.

**Martingale measure:** A martingale measure for a market model is a probability measure  $\mathbb{Q}$  on  $(\Omega, \mathcal{F})$  such that the value of each of the model's assets  $A_i(t)/N(t)$  relative to the numeraire is a martingale, i.e.

$$\frac{A_i(t)}{N(t)} = \mathbb{E}^{\mathbb{Q}} \left[ \frac{A_i(T)}{N(T)} | \mathcal{F}_t \right],$$

where  $\mathbb{E}^{\mathbb{Q}} []$  denotes the expectation with respect to the measure  $\mathbb{Q}$ .

**Equivalent measures:** Two measures on  $(\Omega, \mathcal{F})$  are equivalent if they agree on the possible scenarios for the assets. This means that events with probability 0 according to one measure also have probability 0 according to the other measure.

**derivative instrument:** A derivative in a market model is any financial instrument whose payoff is uniquely determined by the assets.

**self financing strategy :** A self financing strategy (or portfolio) is a dynamically managed portfolio of assets without cash inflows or outflows. More precisely it is a multi-dimensional predictable stochastic process  $(\phi_i(t))$  on the filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$  defining the quantities of assets  $A_i(t)$  held in a portfolio with value  $\sum_i \phi_i(t) A_i(t)$  such that its change in value is solely due to changes in the asset values, i.e.

$$dV(t) = \sum_i \phi_i(t) dA_i(t)$$

This means that you can only change the composition of the dynamic portfolio by selling some assets and investing all its proceeds in buying other assets.

**arbitrage free:** The model is called arbitrage free if prices of derivatives are such that there exists no self-financing portfolio with value  $V(t)$  using the derivative, the assets and the numeraire starting with  $V(0) = 0$  and such that at some time  $t > 0$ ,  $V(t) \geq 0$  and  $P(V(t) > 0) > 0$ . So an arbitrage opportunity is a self-financing portfolio created from scratch that has at some future time  $t$  a guaranteed non-negative value and (strictly) positive value with non-zero probability (i.e. as if it were a free lottery ticket).

**risk neutral measure:** A risk neutral measure with respect to a model is a martingale measure  $\mathbb{Q}$  that is equivalent to the measure  $P$ .

**complete model:** A market model is said to be complete if the payoff of any derivative can be obtained by a replicating strategy using the assets and numeraire. Here a replicating strategy for a derivative is a self-financing portfolio of the assets and numeraire such that its value at payoff date equals the payoff of the derivative.

**First Fundamental Asset Pricing Theorem:** A market model is arbitrage free if and only if there exists a martingale measure equivalent to the real world measure. Such a measure is called a risk neutral measure (for the model). In such a model the value  $V(t)$  of any derivative relative to the numeraire equals its expected value under the risk neutral measure:

$$\frac{V(t)}{N(t)} = \mathbb{E}^{\mathbb{Q}} \left[ \frac{V(T)}{N(T)} | \mathcal{F}_t \right].$$

**Second Fundamental Asset Pricing Theorem:** A market model is complete if and only if there exists exactly one risk neutral measure equivalent to the real world measure.

In practice the assets are often  $T$ -maturity riskfree zero coupon bonds and stocks with current values  $P(0, T)$  and  $S(0)$ . The numeraire is often the risk free bank account. You then set up the model directly in the risk neutral world such that

- It recovers the riskfree zero coupon bond prices

$$\mathbb{E}^{\mathbb{Q}} \left[ e^{-\int_0^T r(s)ds} | \mathcal{F}_0 \right] = P(0, T).$$

where

$$\mathbb{E}^{\mathbb{Q}} []$$

denotes the expectation according to the risk neutral measure,  $r(s)$  denotes the short rate at time  $s$ ,  $\mathcal{F}_0$  denotes the information at the current time 0.

- It recovers the current stock prices

$$\mathbb{E}^{\mathbb{Q}} \left[ S(T) e^{-\int_0^T r(s)ds} | \mathcal{F}_0 \right] = S(0).$$

The condition on stock prices can be fulfilled by using the riskfree rate as drift for the stock price. In case you also need to model risky bonds it is necessary to model spreads as well and such that you recover the current value of the risky bonds. In case you need to model instruments in different currencies you need to recover forward FX rates and foreign riskfree zero coupon bonds as well.

Often one chooses a model from a parametrized family. Some parameters are fixed by the risk neutral conditions above. Other parameters are still free to choose. If possible, they are chosen to match prices of derivatives on the market assets. Otherwise one uses expert opinion or historical information to fix them. Once all parameters are fixed, you can use the model to price any derivative on the market assets.

As an example consider geometric Brownian motion for a stock. This model has two parameters: the drift and volatility. The drift must be taken equal

to the riskfree rate in order to have a risk neutral model. One can then use a plain vanilla call option as extra calibration instrument to fix the volatility. Usually the calibration instrument is chosen such that is as similar as possible to the instrument one finally wants to price. For instance, if one wants to price a lookback call option with a given maturity and strike, one may take the plain vanilla call option with same strike and maturity as calibration instrument.

**Exercise 26.** Consider a discrete time model with two times: 0 and 1 year. As numeraire we take a 1-year unit notional riskfree zero coupon bond  $Z$ . We assume the current value of the bond  $Z(0) = 1$  (i.e. the riskfree rate is 0.) the other asset in the model is a risky asset  $S(t)$  that has start value  $S(0) = 1$  and has two possible values at time 1: either  $S(1) = 2$  or  $S(1) = \frac{1}{2}$ . Each of these events has probability  $\frac{1}{2}$  in the real world. Show that there exists a unique equivalent martingale measure and determine the upward and downward risk neutral probabilities. Next consider a derivative instrument  $V$  on  $S$  such that  $V(1) = 1$  if  $S(1) = 2$  and  $V(1) = 0$  otherwise. Show that  $V(0) = \frac{1}{3}$ . Show how to replicate the derivative by buying and/or short selling the numeraire and the risky asset at time 0.

**Exercise 27.** Consider a similar discrete time model as in the previous exercise, but assume now that  $S(1)$  has three possible values  $2, 1, \frac{1}{2}$  each with real world probability  $\frac{1}{3}$ . Show that there exist several equivalent martingale measures and determine their probabilities. Consider again a derivative  $V$  where  $V(1) = 1$  if  $S(1) = 2$  and  $V(1) = 0$  otherwise. Show that all values  $V(0)$  corresponding to these equivalent martingale measures lie in the interval  $[0, \frac{1}{3}]$ . Show that values in this interval cannot be arbitAGED, but values outside the interval can.

**Exercise 28.** If the numeraire in a market model is a riskfree instrument, show that the expected return (according to an equivalent martingale measure) on any derivative instrument then equals the risk free rate.

## 3.2 Partial Differential Equation for a claim

Let  $x(t)$  denote some financial tradable that follows a SDE

$$dx(t) = \mu(x, t)dt + \sigma(x, t)dB(t),$$

where  $B(t)$  is Brownian motion with respect to a risk neutral measure  $\mathbb{Q}$  (where the numeraire is the bank account process  $\exp(\int_0^t r(u)du)$ .)

Here the drift  $\mu(x, t)$  and volatility  $\sigma(x, t)$  of the process are smooth deterministic functions in two variables.

In the risk neutral world the expected return on a financial tradable must equal the risk free rate  $r(t)$ . Hence we get

$$\mu(x, t) = r(t)x(t)$$

Let  $g(x)$  denote the payoff of a claim on  $x$  at a future time  $T$ . Let  $f(x, t)$  denote the value of the claim at time  $t \leq T$  with  $x(t) = x$ .

By Itô's lemma we have

$$\begin{aligned} df &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dx)^2 \\ &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} r(t) dt + \frac{\partial f}{\partial x} \sigma(x, t) dB(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2(x, t) dt \end{aligned}$$

Taking expectations we get

$$\mathbb{E}^{\mathbb{Q}} [df] = \left( \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} r(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2(x, t) \right) dt$$

Now also the derivative must have the risk free rate as expected return, so

$$\mathbb{E}^{\mathbb{Q}} [df] = r(t) f(x, t) dt$$

and we get an ordinary partial differential equation for  $f$

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \mu(x, t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2(x, t) = r(t) f(x, t)$$

with boundary condition

$$f(x, T) = g(x).$$

This equation is sometimes called the Feynman-Kac equation.

Feynman-Kac proved that the solution of this PDE is

$$f(x, t) = \mathbb{E}^{\mathbb{Q}} \left[ g(x(T)) e^{- \int_t^T r(u) du} | x(t) = x \right],$$

i.e. they retrieve the fundamental asset pricing theorem. For a partial proof of the Feynmac Kac solution without the use of the Fundamental Asset Pricing theorem, see Feynman-Kac Wikipedia

We can use the PDE and its boundary conditions (you will also need to know boundary conditions for extreme values of  $x(t)$ ) to solve the claim value numerically. Note that this gives you an alternative way to value derivatives, apart from the Monte Carlo simulation technique. It is however less general (only works well in low dimensions and does not handle path dependent options).

### 3.3 Partial differential equation for the density

Let  $x(t)$  be as in the previous subsection. Let  $\rho(x, t)$  denote the density function of  $x(t)$  given  $x(0)$ . Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be any function in one variable  $x$  with continuous first and second order partial derivatives for which  $f(x)$  and  $\partial f / \partial x$  go to 0 when  $|x|$  goes to  $\infty$ . According to Itô's lemma

$$\begin{aligned}
df(x(t)) &= \frac{\partial f}{\partial x} dx(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dx(t))^2 \\
&= \frac{\partial f}{\partial x} (\mu(x, t) dt + \sigma(x, t) dB(t)) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2(x, t) dt
\end{aligned}$$

Taking expectations, we get

$$\mathbb{E}^Q [df(x(t))] = \mathbb{E}^Q \left[ \frac{\partial f}{\partial x} \mu(x, t) dt \right] + \frac{1}{2} \mathbb{E}^Q \left[ \frac{\partial^2 f}{\partial x^2} \sigma^2(x, t) dt \right]$$

Hence (using integration by parts)

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x) \frac{\partial \rho}{\partial t}(x, t) dx &= \int_{-\infty}^{\infty} \frac{\partial f}{\partial x} \mu(x, t) \rho(x, t) dx + \frac{1}{2} \int_{-\infty}^{\infty} \frac{\partial^2 f}{\partial x^2} \sigma^2(x, t) \rho(x, t) dx \\
&= - \int_{-\infty}^{\infty} f(x) \frac{\partial(\mu(x, t) \rho(x, t))}{\partial x} dx + \frac{1}{2} \int_{-\infty}^{\infty} f(x) \frac{\partial^2(\sigma^2(x, t) \rho(x, t))}{\partial x^2} dx
\end{aligned}$$

As this equation needs to hold for any function  $f(x)$  satisfying the requirements above,  $\rho(x, t)$  must satisfy the following PDE

$$\frac{\partial \rho}{\partial t} = - \frac{\partial(\mu \rho)}{\partial x} + \frac{1}{2} \frac{\partial(\sigma^2 \rho)}{\partial x^2}$$

This equation is called the Focker-Planck equation.

**Exercise 29.** Explain why

$$\mathbb{E}^Q [df(x(t))] = \int_{-\infty}^{\infty} f(x) \frac{\partial \rho}{\partial t}(x, t) dx$$

### 3.4 Girsanov's Theorem

**Radon-Nikodym derivative :**

Given two equivalent probability measures  $P$  and  $Q$  on a measurable space  $(\Omega, \mathcal{F})$  with filtration  $\mathcal{F}_t$  (i.e. measures with the same null sets). Then we can define a random variable  $\frac{dQ}{dP}$  on  $P$  possible paths such that

$$\mathbb{E}^Q [X_t | \mathcal{F}_s] = \zeta_s^{-1} \mathbb{E}^P [\zeta_t X_t | \mathcal{F}_s]$$

for  $s \leq t$  where

$$\zeta_t = \mathbb{E}^P \left[ \frac{dQ}{dP} | \mathcal{F}_t \right]$$

and  $X_t$  is any process adapted to the filtration  $\mathcal{F}_t$ .

The process  $\zeta_t$  is called the Radon-Nikodym derivative and we often write it in a more sloppy way as  $\frac{dQ}{dP} | \mathcal{F}_t$  or just one of  $\frac{dQ}{dP}(t)$ ,  $\frac{dQ}{dP}(\omega)$ ,  $\frac{dQ}{dP}$ .

**Girsanov's theorem:** Let  $B(t)$  be Brownian motion on a probability space  $(\Omega, \mathcal{F}, P)$ . Let  $X(t)$  be a stochastic process on this space adapted to the natural filtration  $\mathcal{F}_t$  of  $B(t)$ . Define the Doleans-Dade exponential of  $X(t)$  by

$$\mathcal{E}(X)_t = e^{X(t) - \frac{1}{2}[X]_t}.$$

If  $\mathcal{E}(X)_t$  is a strictly positive martingale, a probability measure  $Q$  on  $(\Omega, \mathcal{F})$  can be defined such that its Radon-Nikodym derivative equals

$$\frac{dQ}{dP} | \mathcal{F}_t = \mathcal{E}(X)_t$$

and  $Q$  is equivalent to  $P$  on each  $\mathcal{F}_t$ . The opposite also holds. Furthermore, if  $Y_t^P$  is a  $P$ -local martingale on  $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ , then

$$Y_t^Q = Y_t^P - [X, Y_t^P]_t$$

is a  $Q$ -local martingale on  $(\Omega, \mathcal{F}, Q, \mathcal{F}_t)$ . In the special case that  $Y_t^P$  is a  $P$ -Brownian motion and  $X_t$  is continuous,  $Y_t^Q$  is  $Q$ -Brownian motion.

This theorem is often applied to the case when

$$X_t = \int_0^t \alpha(s) dB(s).$$

Here  $\alpha$  does not need to be a deterministic function of time (it can be a stochastic process as long as the integral can be defined). In the special case that  $(B, Y_t^P)$  is joint  $P$ -Brownian motion with correlation  $\rho(t)$  and  $\alpha(s)$  is deterministic then we have

$$dY^Q(t) + \alpha(t)\rho(t)dt = dY^P(t).$$

In other words, under  $Q$  the process  $Y^P(t)$  is Brownian motion plus drift, where the drift coefficient is given by  $\alpha(t)\rho(t)$ .

**Exercise 30.** Show that  $[X]_t = \int_0^t \alpha(s)^2 ds$  when  $X_t = \int_0^t \alpha(s) dB(s)$  and show that then  $\mathcal{E}(X)_t$  is indeed a martingale.

**Exercise 31.** Let  $B(t)$  be a Brownian motion on  $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ . Show directly that  $B(t) - \mu t$  is Brownian motion on  $(\Omega, \mathcal{F}, Q, \mathcal{F}_t)$  where

$$\frac{dQ}{dP} | \mathcal{F}_t = e^{\mu B(t) - \frac{\mu^2 t}{2}}.$$

**Exercise 32.** Let  $r_d(s)$  and  $r_f(s)$  denote the riskfree short rate in a domestic and foreign currency. Let  $X(t)$  denote the value of one unit of the foreign currency expressed in units of the domestic currency. Let

$$B_d(t) = e^{\int_0^t r_d(s) ds}$$

and

$$B_f(t) = e^{\int_0^t r_f(s) ds}$$

denote the corresponding riskfree bank accounts. Let  $P_d$  and  $P_f$  denote risk neutral measures associated to these bank account numeraires. Assume  $B_d(0) = B_f(0) = X(0) = 1$  for sake of exposition. Assume

$$\frac{dX(t)}{X(t)} = (r_d(t) - r_f(t))dt + \sigma_X dW^{P_d}(t)$$

where  $W^{P_d}(t)$  is  $P_d$ -Brownian motion. Let  $Y(t) = X(t)B_f(t)/B_d(t)$ . Let  $S(t)$  denote the value of a foreign asset in units of the foreign currency. Assume that

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma_S dB^{P_f}(t)$$

with  $[B^{P_f}(t), W^{P_d}(t)] = \rho t$  where  $B^{P_f}(t)$  is  $P_f$ -Brownian motion.

Show that

1.

$$\frac{dP_f}{dP_d} | \mathcal{F}_t = Y(t)$$

2.

$$Y(t) = \mathcal{E}(\sigma_X W^P(t))$$

3.

$$B^{P_f}(t) = B^{P_d}(t) - \sigma_X \rho t$$

where  $B^{P_d}(t)$  is  $P_d$ -Brownian motion.

4.

$$\frac{dS(t)}{S(t)} = (\mu - \rho \sigma_X \sigma_S)dt + \sigma_S dB^{P_d}(t)$$

**Exercise 33.** Let  $r_d(s)$  and  $r_f(s)$  denote the riskfree short rate in a domestic and foreign currency. Let  $X(t)$  denote the value of one unit of the foreign currency expressed in units of the domestic currency. Let

$$B_d(t) = e^{\int_0^t r_d(s)ds}$$

and

$$B_f(t) = e^{\int_0^t r_f(s)ds}$$

denote the corresponding riskfree bank accounts. Let  $P_d$  and  $P_f$  denote risk neutral measures associated to these bank account numeraires. Assume

$$\frac{dX(t)}{X(t)} = \mu_X(t)dt + \sigma_X dW_X^{P_d}(t)$$

where  $W_X^{P_d}(t)$  is  $P_d$ -Brownian motion. Let  $Y(t) = X(t)B_f(t)/B_d(t)$ . Let  $S(t)$  denote the value of a foreign tradable asset in units of the foreign currency. Assume that

$$\frac{dS(t)}{S(t)} = \mu_S(t)dt + \sigma_S dW_S^{P_f}(t)$$

with  $[W_S^{P_f}(t), W_X^{P_d}(t)] = \rho t$  where  $W_S^{P_f}(t)$  is  $P_f$ -Brownian motion.

Show that

1.

$$\frac{dB_f(t)}{B_f(t)} = r_f(t)dt$$

2.

$$\frac{d(S(t)/B_f(t))}{S(t)/B_f(t)} = \frac{dS(t)}{S(t)} - \frac{dB_f(t)}{B_f(t)}$$

3.

$$\mu_S(t) = r_f(t)$$

4.

$$\frac{d(X(t)B_f(t)/B_d(t))}{X(t)B_f(t)/B_d(t)} = \frac{dX(t)}{X(t)} + \frac{dB_f(t)}{B_f(t)} - \frac{dB_d(t)}{B_d(t)}$$

5.

$$\mu_X(t) = r_d(t) - r_f(t)$$

6.

$$\frac{d(S(t)X(t))}{S(t)X(t)} = \frac{dS(t)}{S(t)} + \frac{dX(t)}{X(t)} + \frac{dS(t)}{S(t)} \frac{dX(t)}{X(t)}$$

7.

$$\frac{d(S(t)X(t))}{S(t)X(t)} = r_d(t)dt + \sigma_S W_S^{P_d}(t) + \sigma_X dW_X^{P_d}(t)$$

for a  $P_d$ -Brownian motion

$$W_S^{P_d}(t) = W_S^{P_f}(t) + \rho\sigma_X t$$

8.

$$\frac{d(S(t)X(t))}{S(t)X(t)} = r_d(t)dt + \sigma_{SX} W_{SX}^{P_d}(t)$$

for a  $P_d$ -Brownian motion

$$W_{SX}^{P_d}(t)$$

and  $\sigma_{SX}^2 = \sigma_X^2 + \sigma_S^2 + 2\rho\sigma_X\sigma_S$

A multivariate extension of Girsanov's theorem leads to the following result that can be used as a toolkit for changing numeraires in case of Brownian motions.

**Change of numeraire toolkit:** Let  $X(t) = (X_i(t))$  be an  $n$ -dimensional process satisfying

$$\frac{dX_i(t)}{X_i(t)} = \mu_i^S(t)dt + \sigma_i(t)C_idB^S(t),$$

where the volatility coefficients  $\sigma_i(t)$  are deterministic functions of time, the  $C_i$  are  $1 \times n$  vectors and  $B^S(t)$  is  $n \times 1$  standard Brownian motion under the measure associated to a numeraire  $S$ . Here the matrix  $C$  with the  $C_i$  as rows is used to induce a correlation matrix  $\rho = CC'$ . Under another numeraire  $U$  we then have (by Girsanov's theorem)

$$\frac{dX_i(t)}{X_i(t)} = \mu_i^U(t)dt + \sigma_i(t)C_idB^U(t)$$

for a standard Brownian motion  $B^U(t)$  under the measure associated to  $U$ . Then

$$d\mu_i^U(t) = d\mu_i^S(t) - d\langle \ln X_i(t), \ln(S/U) \rangle.$$

Note that the drift coefficients  $\mu_i^S(t)$  and  $\mu_i^U(t)$  may be stochastic.

### 3.5 Derivative pricing in multi-curve paradigm

Up to now we tacitly assumed a risk free rate  $r(t)$  at which banks can operate. In reality traders pay different rates depending the kind of transactions they do in trading and hedging derivatives. Piterbarg [Pit10] considers three different rates playing a role:

- $r_C(t)$  : interest rate for funding under CSA agreement, i.e. on almost risk free collateral
- $r_R(t)$  : interest rate for stock funding, the so-called repo rate for borrowing stocks with the stock as collateral
- $r_F(t)$  : interest rate for unsecured funding, i.e. a rate incorporating the bank's credit risk

Piterbarg then considers an asset  $x(t)$  following

$$dx(t)/x(t) = \mu_x(t)dt + \sigma_x(t)dB(t)$$

for a Brownian motion under the real measure  $P$ . He assumes the asset has dividend rate  $r_D(t)$  and derives an SDE for the value  $f(x, t)$  of a derivative on the asset by replication arguments:

$$\frac{\partial f}{\partial t} + (r_R(t) - r_D(t))\frac{\partial f}{\partial x} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2(x, t) = r_F(t)f(x, t) - (r_F(t) - r_C(t))C(t)$$

where  $C(t)$  is the value of the posted collateral.

This can be solved by an extension of Feynman-Kac to yield

$$f(x, t) = \mathbb{E}^{\mathbb{Q}} \left[ f(x(T), T) e^{-\int_t^T r_F(u) du} + \int_t^T e^{-\int_t^u r_F(v) dv} (r_f(u) - r_C(u)) C(u) du | x(t) = x \right],$$

in a measure  $\mathbb{Q}$  in which the asset grows at rate  $r_R(t) - r_D(t)$ , i.e.

$$dx(t)/x(t) = (r_R(t) - r_D(t))dt + \sigma_x(t)dB_x(t)$$

where  $B_x(t)$  is  $\mathbb{Q}$ -Brownian motion. This means that we should replace the classical risk free rate by the repo rate in modelling the asset.

Equivalently we have the formula

$$\begin{aligned} f(x, t) &= \mathbb{E}^{\mathbb{Q}} \left[ f(x(T), T) e^{-\int_t^T r_C(u) du} | x(t) = x \right] \\ &\quad - \mathbb{E}^{\mathbb{Q}} \left[ \int_t^T e^{-\int_t^u r_C(v) dv} (r_F(u) - r_C(u)) (f(x(u), u) - C(u)) du | x(t) = x \right] \end{aligned}$$

Without collateral we get

$$f(x, t) = \mathbb{E}^{\mathbb{Q}} \left[ f(x(T), T) e^{-\int_t^T r_F(u) du} | x(t) = x \right]$$

So, in that case we should replace the classical risk free rate by the bank's funding rate in discounting.

With full collateral we get

$$f(x, t) = \mathbb{E}^{\mathbb{Q}} \left[ f(x(T), T) e^{-\int_t^T r_C(u) du} | x(t) = x \right]$$

So, in that case we should replace the classical risk free rate by the collateral rate in discounting.

# Chapter 4

## Basic models

### 4.1 Geometric Brownian motion

Geometric Brownian motion  $S(t)$  is usually introduced via a SDE (Stochastic differential equation):

$$d(S(t)) = \mu(t)S(t)dt + \sigma(t)S(t)dB(t).$$

This actually means that  $S(t)$  must solve the stochastic Itô integral equation

$$S(t) = \int_0^t \mu(s)S(s)ds + \int_0^t \sigma(s)S(s)dB(s).$$

Here  $\mu(t)$  and  $\sigma(t)$  are assumed to be deterministic functions of time. Defining the logarithmic return process  $X(t) = \ln(\frac{S(t)}{S(0)})$  we get (using Ito's lemma)

$$\begin{aligned} dX(t) &= \frac{S(0)}{S(t)}d\left(\frac{S(t)}{S(0)}\right) - \frac{1}{2}\frac{S^2(0)}{S^2(t)}\left(d\left(\frac{S(t)}{S(0)}\right)\right)^2 \\ &= (\mu(t) - \frac{1}{2}\sigma^2(t))dt + \sigma(t)dB(t) \end{aligned}$$

It follows that

$$X(t) = \int_0^t (\mu(s) - \frac{1}{2}\sigma^2(s))ds + \int_0^t \sigma(s)dB(s)$$

and

$$S(t) = S(0)\exp\left(\int_0^t (\mu(s) - \frac{1}{2}\sigma^2(s))ds + \int_0^t \sigma(s)dB(s)\right).$$

This model leads to the famous Black-Scholes formula for pricing call (or put) options (see [BS73]). Indeed, we can immediately assume that  $B(t)$  is

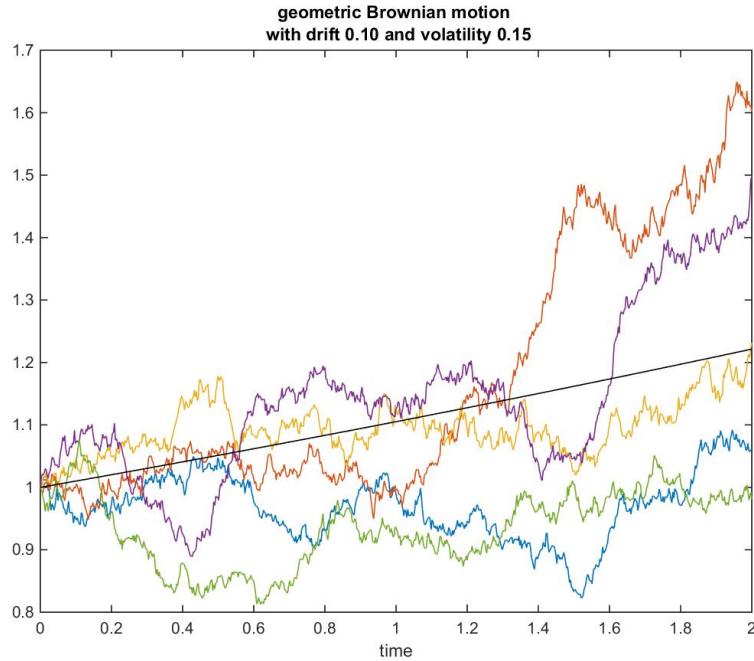


Figure 4.1: 5 sample paths of geometric Brownian motion

Brownian motion with respect to the risk neutral measure, i.e. the measure  $Q$  associated to the bank account with value

$$C(t) = e^{\int_0^t r(s)ds},$$

where  $r(s)$  denotes the risk free short rate (assumed to be deterministic).

**Exercise 34.** Show that

$$\int_0^t \sigma(s) dB(s) \sim N(0, \int_0^t \sigma^2(s) ds).$$

**Exercise 35.** Show that

$$\mathbb{E}^Q \left[ \exp \left( -\frac{1}{2} \int_0^t \sigma^2(s) ds + \int_0^t \sigma(s) B(s) \right) \right] = 1.$$

Hint: show that

$$E(e^X) = e^{\mu + \frac{1}{2}\sigma^2}$$

for a  $N(\mu, \sigma^2)$  distributed normal random variable  $X$ .

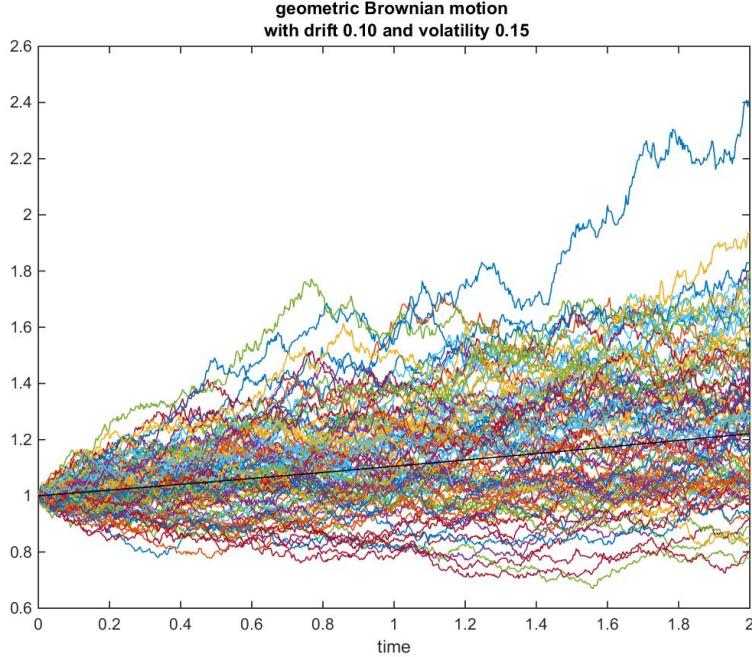


Figure 4.2: 100 sample paths of geometric Brownian motion

Then we must have

$$S(0) = \mathbb{E}^Q [S(t)/C(t)] = S(0)e^{\int_0^t (\mu(s) - r(s))ds},$$

and hence the drift  $\mu(t) = r(t)$ . The payoff of a call option with strike  $K$  maturing at time  $T$  equals

$$\max(S(T) - K, 0),$$

such that the value of the call at time 0 equals

$$\begin{aligned} V(0) &= \mathbb{E}^Q [\max(S(T) - K, 0)/C(T)] \\ &= e^{-\int_0^T r(t)dt} \mathbb{E}^Q \left[ \max(S(0)e^{\int_0^T (r(t) - \frac{1}{2}\sigma^2(t))dt + \int_0^T \sigma(s)dB(s)} - K, 0) \right] \\ &= e^{-\int_0^T r(t)dt} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \max(S(0)e^{\int_0^T r(t)dt - \frac{1}{2}\Sigma^2(T)T + x\Sigma(T)\sqrt{T}} - K, 0) e^{-\frac{x^2}{2}} dx \\ &= S(0)N(d_+) - e^{-\int_0^T r(t)dt} KN(d_-) \end{aligned}$$

where

$$d_{\pm} = \frac{\ln(S(0)/K) + \int_0^T r(t)dt \pm T\Sigma^2(T)/2}{\Sigma(T)\sqrt{T}},$$

$$T\Sigma^2(T) = \int_0^T \sigma^2(s)ds$$

and  $N$  denotes the standard normal distribution function. Here  $\Sigma(T)$  is often called the Black-Scholes volatility.

The value of  $\Sigma(T)$  corresponding to the market price of a call option with strike  $T$  and maturity  $T$  is often called the Black-Scholes implied volatility and denoted by  $\Sigma(T, K)$ .

## 4.2 Vasicek's short rate model

We refer to [BM06] for a detailed analysis of Vasicek's model or to the original paper of Vasicek [Vas76].

Vasicek's model is a standard choice to model interest rates. It is expressed in terms of the short rate

$$r(t) = \lim_{dt \rightarrow 0} R(t, t + dt)$$

where  $R(t, T)$  denotes the continuously compounded zero coupon rate for maturity  $T$  at time  $t$ . The model is then defined by the SDE

$$dr(t) = \kappa(\theta - r(t))dt + \sigma dB(t)$$

where  $B(t)$  denotes standard Brownian motion.

The model has the following features

- Analytic formulas for bond prices and bond option prices in terms of short rate
- Exact simulation of short rate and integrated short rate possible

The first feature allows us to calibrate the model parameters to bond prices and bond option prices. Due to the fact that the model has only three parameters, the calibration will be far from perfect though.

The second feature (combined with the first feature) allows us to value most interest rate derivatives by Monte Carlo simulation. The integrated short rate is needed to determine the discount factor to be applied to future cash flows, whereas the short rate itself can be used to determine the interest rate curve at any future time step. So, you can value any derivative whose payoffs at times  $t$  are determined by the evolution of the interest rate curve up to time  $t$ .

Taking expectations we get

$$d\mathbb{E}[r(t)|r(s)] = \kappa(\theta - \mathbb{E}[r(t)|r(s)])dt,$$

and we may solve

$$\mathbb{E}[r(t)|r(s)] = \theta + (r(s) - \theta)e^{-\kappa(t-s)}.$$

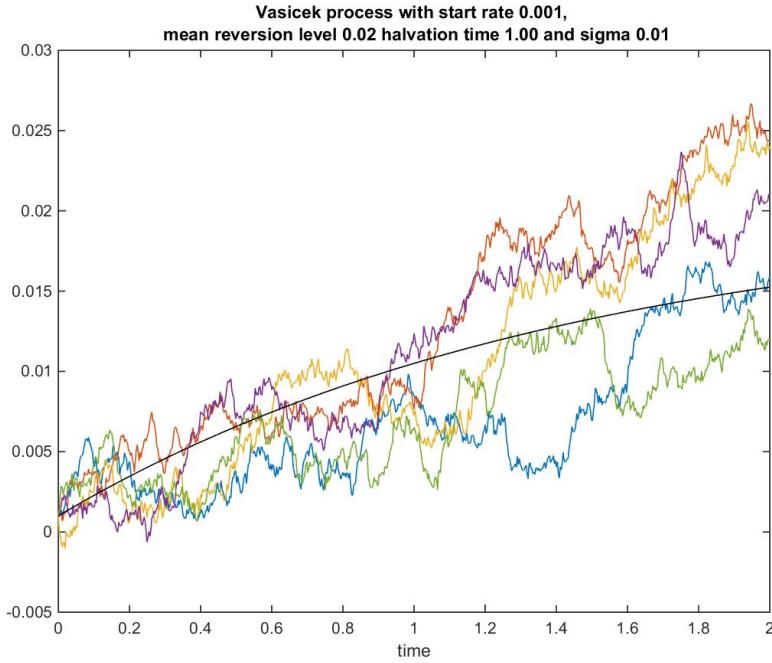


Figure 4.3: 5 sample paths of a Vasicek process

This clearly shows that on average the process  $r(t)$  reverts to its mean level  $\theta$  when  $t \rightarrow \infty$ . Hence  $\theta$  is called the mean reversion level and  $\kappa$  is called the mean reversion speed. We also define the halvation time  $\tau$  by setting

$$e^{-\kappa\tau} = \frac{1}{2}.$$

To solve the SDE set  $x(t) = r(t) - \theta$ . The SDE can then be rewritten as

$$dx(t) = -\kappa x(t)dt + \sigma dB(t).$$

Using Itô's lemma we get

$$\begin{aligned} d(x(t)e^{\kappa t}) &= \kappa e^{\kappa t}x(t)dt + e^{\kappa t}dx(t) \\ &= \sigma e^{\kappa t}dB(t) \end{aligned}$$

So

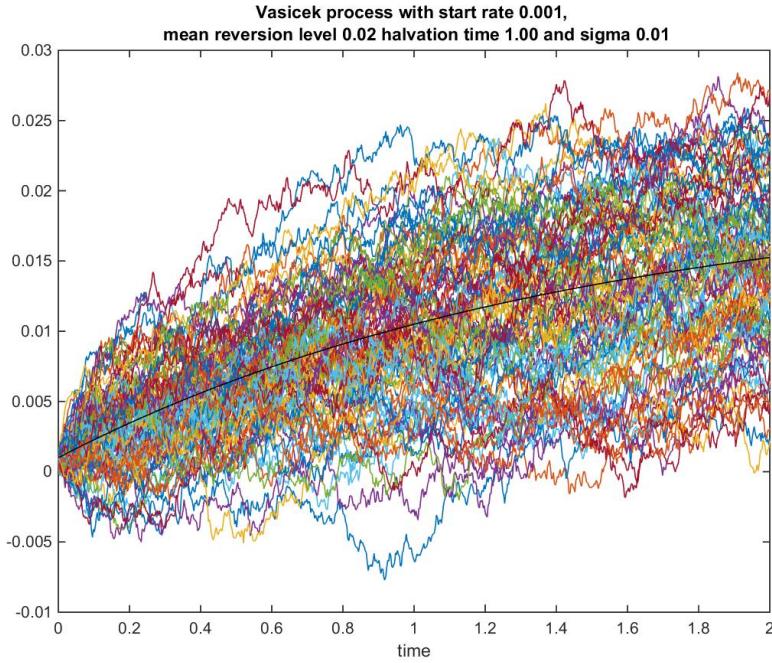


Figure 4.4: 100 sample paths of a Vasicek process

$$\begin{aligned}
 x(t)e^{\kappa t} &= x(s)e^{\kappa s} + \sigma \int_s^t e^{\kappa u} dB(u) \\
 x(t) &= x(s)e^{-\kappa(t-s)} + \sigma \int_s^t e^{-\kappa(t-u)} dB(u) \\
 r(t) &= \theta + (r(s) - \theta)e^{-\kappa(t-s)} + \sigma \int_s^t e^{-\kappa(t-u)} dB(u) \\
 &= r(s)e^{-\kappa(t-s)} + \theta(1 - e^{-\kappa(t-s)}) + \sigma \int_s^t e^{-\kappa(t-u)} dB(u)
 \end{aligned}$$

From the Itô isometry it now follows that

$$r(t)|r(s) \sim N\left(\theta + (r(s) - \theta)e^{-\kappa(t-s)}, \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa(t-s)})\right).$$

Integrating  $r(u)$  over  $[t, T]$  we get

$$\int_t^T r(u) du = r(t)B(t, T) + \kappa\theta \int_t^T B(t, u) du + \sigma \int_t^T B(u, T) dB(u),$$

where

$$B(t, T) = \int_t^T e^{-\kappa(u-t)} du = \frac{1 - e^{-\kappa(T-t)}}{\kappa}.$$

It follows that

$$\int_t^T r(u) du | r(t) \sim N(E(t, T), V(t, T)),$$

where

$$\begin{aligned} E(t, T) &= r(t)B(t, T) + \theta(T - t - B(t, T)) \\ V(t, T) &= \sigma^2 \int_t^T B(u, T)^2 du \\ &= \frac{\sigma^2}{\kappa^2} \left( T - t + \frac{2}{\kappa} e^{-\kappa(T-t)} - \frac{e^{-2\kappa(T-t)}}{2\kappa} - \frac{3}{2\kappa} \right) \\ &= \frac{\sigma^2}{\kappa^2} \left( T - t - B(t, T) - \frac{\kappa}{2} B(t, T)^2 \right) \end{aligned}$$

It is clear also the joint distribution of  $r(T)$  and  $\int_t^T r(s)ds$  conditional on  $r(t)$  is a normal distribution and one can obtain analytic expressions for its covariance matrix. This enables us to simulate both the short rate and its integral over any time interval exactly (without need to discretise).

**Exercise 36.** Explain why the joint distribution of  $r(T)$  and  $\int_0^T r(s)ds$  in vasicek's model is a bivariate normal distribution and compute the covariance

$$\text{Cov}(r(T), \int_0^T r(s)ds) = \frac{\sigma^2}{\kappa^2} \left( \frac{1}{2} + \frac{1}{2} e^{-2\kappa T} - e^{-\kappa T} \right).$$

The zero coupon bond price at time  $t$  is given by

$$\begin{aligned} P(t, T) &= \mathbb{E} \left[ e^{-\int_t^T r(s)ds} | r(t) \right] \\ &= e^{-E(t, T) + \frac{V(t, T)}{2}} \\ &= A(t, T) e^{-B(t, T)r(t)} \end{aligned}$$

where

$$A(t, T) = \exp \left\{ \left( \theta - \frac{\sigma^2}{2\kappa^2} \right) (B(t, T) - T + t) - \frac{\sigma^2}{4\kappa} B(t, T)^2 \right\}$$

This means that we can easily derive the complete yield curve at any future time from the short rate at that time.

We may also obtain the zero coupon bond values more directly by using the Feynman-Kac equation:

$$r(t)P(t, T) = \frac{\partial P}{\partial t} + \frac{\partial P}{\partial r}\kappa(\theta - r) + \frac{1}{2} \frac{\partial^2 P}{\partial r^2}\sigma^2$$

Assuming that  $P(t, T)$  can be written as

$$P(t, T) = A(t, T)e^{-B(t, T)r(t)}$$

for smooth functions  $A(t, T)$  and  $B(t, T)$ , we get

$$r(t)P(t, T) = \frac{\partial A(t, T)/\partial t}{A(t, T)}P(t, T) - \frac{\partial B(t, T)}{\partial t}r(t)P(t, T) - B(t, T)P(t, T)\kappa(\theta - r(t)) + \frac{1}{2}\sigma^2B(t, T)^2P(t, T)$$

Since this equation needs to hold for all  $r(t)$  we get a system of ordinary PDEs:

$$\begin{aligned} 1 &= -\frac{\partial B(t, T)}{\partial t} + B(t, T)\kappa \\ 0 &= \frac{\partial A(t, T)/\partial t}{A(t, T)} - B(t, T)\kappa\theta + \frac{1}{2}\sigma^2B(t, T)^2 \end{aligned}$$

These equations can easily be solved by standard techniques and of course their solution is the same as above.

### 4.3 Hull-White's model

We refer to [BM06] or the original paper of Hull and White [HW90] for more information.

In Vasicek's model the zero coupon bond prices at the current time 0 do not need to agree with the observed zero coupon bond prices in the market. Hull-White's model is an extension of Vasicek's model where the constant  $\theta$  is replaced by a deterministic function  $\theta(t)$ . This function is chosen such that we recover the market's zero coupon bond prices at time 0. The model is specified by the SDE

$$dr(t) = \kappa(\theta(t) - r(t))dt + \sigma dB(t)$$

where

$$\theta(t) = f^M(0, t) + \frac{1}{\kappa} \frac{\partial f^M(0, t)}{\partial t} + \frac{\sigma^2}{2\kappa^2} (1 - e^{-2\kappa t})$$

Here

$$f^M(0, t) = -\frac{\partial \ln P^M(0, t)}{\partial t}$$

denotes the market's current instantaneous forward rates and

$$P^M(0, t) = e^{-\int_0^t f^M(0, u) du}$$

denote the market's zero coupon bond prices. It is assumed that the full curve is known. Usually this curve is obtained by interpolating observed zero coupon rates. The interpolated zero-coupon rate curve must be smooth in order to guarantee the existence of the first derivatives.

The Hull-White model has the same attractive features as Vasicek's model plus market consistency with bond prices:

The model has the following features

- Analytic formulas for bond prices and bond option prices in terms of short rate
- Exact simulation of short rate and integrated short rate possible
- Exact valuation of the market's bond prices

We now derive the formulae. We start by following the same lines as for Vasicek's model but assuming  $\theta(t)$  is any smooth deterministic function. Later we will fix  $\theta(t)$  to match the market's zero coupon bond prices.

Setting  $x(t) = r(t) - \theta(t)$  we can use Itô's lemma to obtain

$$x(t) = x(s)e^{-\kappa(t-s)} - \int_s^t e^{-\kappa(t-u)} d\theta(u) + \sigma \int_s^t e^{-\kappa(t-u)} dB(u)$$

By substituting  $r(t) - \theta(t)$  for  $x(t)$  and using partial integration we get

$$r(t) = r(s)e^{-\kappa(t-s)} + \kappa \int_s^t \theta(u)e^{-\kappa(t-u)} du + \sigma \int_s^t e^{-\kappa(t-u)} dB(u)$$

Integrating the short rate over  $[t, T]$  we get

$$\int_t^T r(u) du = r(t)B(t, T) + \kappa \int_t^T \theta(u)B(u, T) du + \sigma \int_t^T B(u, T) dB(u)$$

where

$$B(t, T) = \int_t^T e^{-\kappa(T-u)} du = \frac{1 - e^{-\kappa(T-t)}}{\kappa}.$$

It follows that

$$\int_t^T r(u) du | r(t) \sim N(E(t, T), V(t, T)),$$

where

$$\begin{aligned}
E(t, T) &= r(t)B(t, T) + \kappa \int_t^T \theta(u)B(u, T)du \\
V(t, T) &= \sigma^2 \int_t^T B(u, T)^2 du \\
&= \frac{\sigma^2}{\kappa^2} \left( T - t + \frac{2}{\kappa} e^{-\kappa(T-t)} - \frac{e^{-2\kappa(T-t)}}{2\kappa} - \frac{3}{2\kappa} \right) \\
&= \frac{\sigma^2}{\kappa^2} \left( T - t - B(t, T) - \frac{\kappa}{2} B(t, T)^2 \right)
\end{aligned}$$

The zero coupon bond prices at time  $t$  are given by

$$\begin{aligned}
P(t, T) &= \mathbb{E} \left[ e^{-\int_t^T r(s)ds} | r(t) \right] \\
&= e^{-E(t, T) + \frac{V(t, T)}{2}}
\end{aligned}$$

We can now use this formula to determine  $\theta(t)$  such that the model reproduces the market's zero coupon bond prices at time 0. Let

$$f^M(0, t) = -\frac{\partial \ln P^M(0, t)}{\partial t}$$

denote the market's current instantaneous forward rates. Then

$$f^M(0, t) = -\frac{\partial \ln P^M(0, t)}{\partial t} \quad (4.1)$$

$$= \frac{\partial E(0, t)}{\partial t} - \frac{1}{2} \frac{\partial V(0, t)}{\partial t} \quad (4.2)$$

$$= r(0)e^{-\kappa t} + \kappa e^{-\kappa t} \int_0^t \theta(u)e^{\kappa u} du - \frac{1}{2} \frac{\partial V(0, t)}{\partial t} \quad (4.3)$$

Taking the partial derivative with respect to  $t$  we get

$$\frac{\partial f^M(0, t)}{\partial t} = -\kappa \left( f^M(0, t) + \frac{1}{2} \frac{\partial V(0, t)}{\partial t} \right) + \theta(t)\kappa - \frac{1}{2} \frac{\partial^2 V(0, t)}{\partial t^2}$$

Using

$$\begin{aligned}
\frac{\partial V(0, t)}{\partial t} &= \frac{\sigma^2}{\kappa^2} (1 - 2e^{-\kappa t} + e^{-2\kappa t}) \\
\frac{\partial^2 V(0, t)}{\partial t^2} &= \frac{\sigma^2}{\kappa^2} (2\kappa e^{-\kappa t} - 2\kappa e^{-2\kappa t})
\end{aligned}$$

we get

$$\theta(t) = f^M(0, t) + \frac{1}{\kappa} \frac{\partial f^M(0, t)}{\partial t} + \frac{\sigma^2}{2\kappa^2} (1 - e^{-2\kappa t})$$

Setting

$$\alpha(t) = f^M(0, t) + \frac{1}{2} \frac{\partial V(0, t)}{\partial t}$$

from 4.3 we get

$$\begin{aligned} \kappa \int_t^T \theta(u) e^{-\kappa(T-u)} du &= \kappa \int_0^T \theta(u) e^{-\kappa(T-u)} du - \kappa \int_0^t \theta(u) e^{-\kappa(T-u)} du \\ &= \alpha(T) + r(0)e^{-\kappa T} - e^{-\kappa(T-t)}(\alpha(t) + r(0)e^{-\kappa t}) \\ &= \alpha(T) - e^{-\kappa(T-t)}\alpha(t) \end{aligned}$$

On the other hand

$$\begin{aligned} \int_t^T \theta(u) du &= \int_t^T \left( f^M(0, u) + \frac{1}{\kappa} \frac{\partial f^M(0, u)}{\partial t} + \frac{\sigma^2}{2\kappa^2} (1 - e^{-2\kappa u}) \right) du \\ &= \ln \frac{P^M(0, t)}{P^M(0, T)} + \frac{1}{\kappa} (f^M(0, T) - f^M(0, t)) + \frac{\sigma^2}{2\kappa^2} \int_t^T (1 - e^{-2\kappa u}) du \end{aligned}$$

Hence

$$\begin{aligned} E(t, T) &= B(t, T)r(t) + \int_t^T \theta(u) du - \int_t^T \theta(u) e^{-\kappa(T-u)} du \\ &= B(t, T)(r(t) - \alpha(t)) + \frac{\alpha(t) - \alpha(T)}{\kappa} + \int_t^T \theta(u) du \\ &= B(t, T)(r(t) - \alpha(t)) + \frac{f^M(0, t) - f^M(0, T)}{\kappa} + \frac{1}{2\kappa} \left( \frac{\partial V(0, t)}{\partial t} - \frac{\partial V(0, T)}{\partial T} \right) + \int_t^T \theta(u) du \\ &= B(t, T)(r(t) - \alpha(t)) + \frac{1}{2\kappa} \left( \frac{\partial V(0, t)}{\partial t} - \frac{\partial V(0, T)}{\partial T} \right) + \ln \frac{P^M(0, t)}{P^M(0, T)} + \frac{\sigma^2}{2\kappa^2} \int_t^T (1 - e^{-2\kappa u}) du \\ &= B(t, T)(r(t) - \alpha(t)) + \ln \frac{P^M(0, t)}{P^M(0, T)} + \frac{1}{2} (V(0, T) - V(0, t)) \end{aligned}$$

We can check that

$$V(0, T) - V(0, t) - V(t, T) - B(t, T) \frac{\partial V(0, t)}{\partial t} = \frac{\sigma^2}{2\kappa} B(t, T)^2 (1 - e^{-2\kappa t}).$$

Using this formula it follows that

$$P(t, T) = A(t, T) e^{-B(t, T)r(t)}$$

where

$$A(t, T) = \frac{P^M(0, T)}{P^M(0, t)} \exp \left( B(t, T) f^M(0, t) - \frac{\sigma^2}{4\kappa} B(t, T)^2 (1 - e^{-2\kappa t}) \right).$$

## 4.4 CIR model

We refer to [BM06] or the original paper [CIR85] for more details.

The Cox-Ingwersen-Ross (CIR) model is a short rate model similar to Vasicek's model. However the volatility is now proportional to the square root of the rate. It follows that under suitable conditions for the parameters the short rates remain positive. It is described by the SDE

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{r(t)}dB(t).$$

The mean reversion part is the same as in Vasicek's model. It can be shown that short rates remain positive under the condition that  $2\kappa\theta > \sigma^2$ .

The CIR model combines the following features

- Analytic formulas for bond prices and bond option prices in terms of short rate
- Exact simulation of the short rate is possible
- Positive short rates with positively skewed distribution

One can slightly extend the model allowing for negative but floored short rates by applying a constant shift. Such model is specified by

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{r(t) - f}dB(t),$$

where  $f$  denotes the floor. Short rates are then larger than the floor  $f$  that can be chosen negative.

**Exercise 37.** Consider the lognormal short model with mean reversion specified by the SDE

$$dr(t) = \kappa(\theta - r(t))dt + \sigma(r(t) - \theta)dB(t).$$

Show that

$$E(e^{\int_0^t r(s)ds}) = \infty.$$

In other words the bank account is expected to explode: for this reason the model is not realistic. *Hint:* Approximate  $\int_0^{\Delta t} r(s)ds$  by  $\frac{r(0)+r(\Delta t)}{2}$  for a small time step  $\Delta t$ .

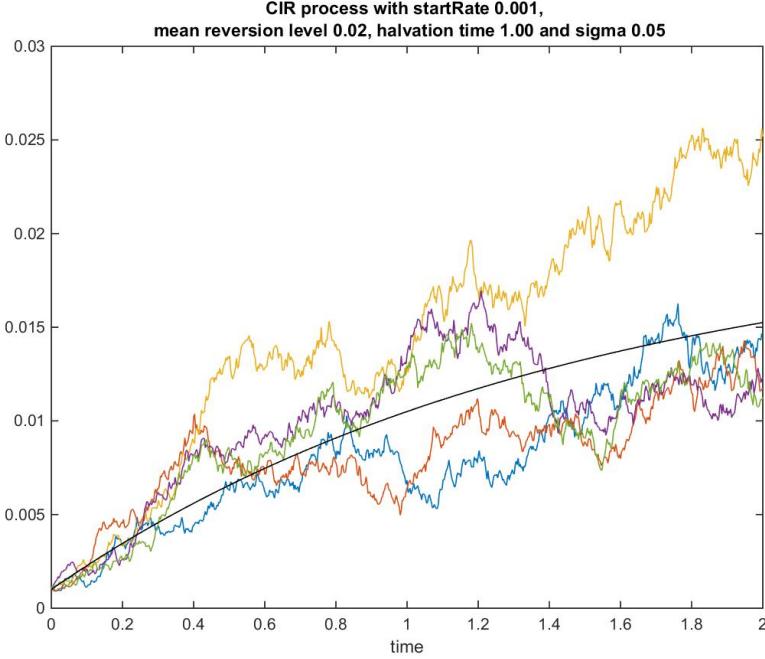


Figure 4.5: 5 sample paths of a CIR process

The probability density function of the short rate  $r(T)$  conditional on its value  $r(t)$  at time  $t$  is given by

$$\rho(r(T)|r(t)) = de^{-u-v} \left(\frac{v}{u}\right)^{\frac{q}{2}} I_q(2\sqrt{uv})$$

where

$$\begin{aligned} d &= \frac{2\kappa}{\sigma^2(1 - e^{-\kappa(T-t)})} \\ u &= dr(t)e^{-\kappa(T-t)} \\ v &= dr(T) \\ q &= \frac{2\kappa\theta}{\sigma^2} - 1 \end{aligned}$$

and  $I_q(\cdot)$  is the modified Bessel function of the first kind of order  $q$ .

The distribution function is the noncentral chi-square,  $\chi^2(2dr(T); 2q+2, 2u)$  with  $2q+2$  degrees of freedom and parameter of noncentrality  $2u$  proportional to the spot rate. We refer to the appendices for a derivation of the formula.

The prices of zero coupon bonds  $P(t, T)$  can be obtained by solving the Feynman-Kac equation for the bond price. As in Vasicek's model we assume an

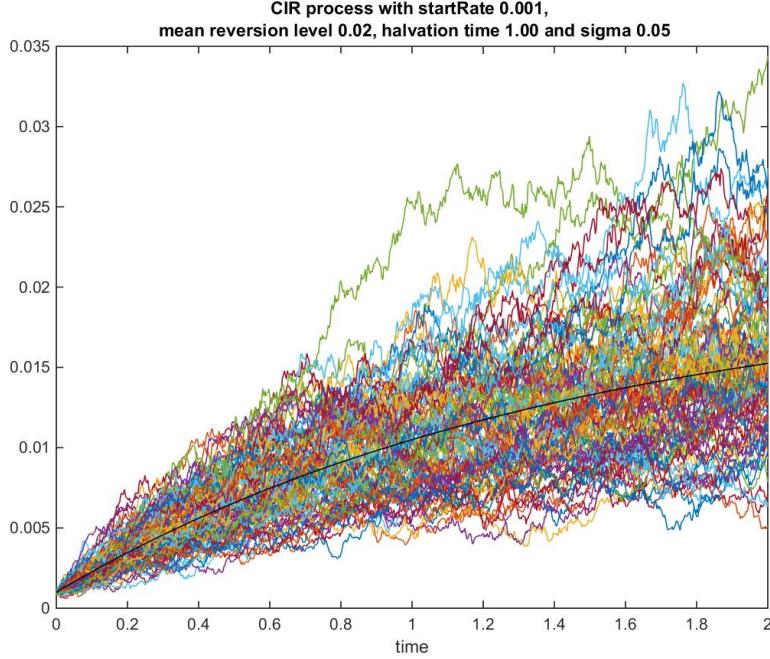


Figure 4.6: 100 sample paths of a CIR process

affine bond price structure , i.e.

$$P(t, T) = A(t, T)e^{-B(t, T)r(t)}$$

for smooth functions  $A(t, T)$  and  $B(t, T)$ .

See appendix 9.2 for a derivation of the solution

$$\begin{aligned} B(t, T) &= \frac{2(e^{\psi(T-t)} - 1)}{(\psi + \kappa)e^{\psi(T-t)} + \psi - \kappa} \\ A(t, T) &= \left( \frac{2\psi e^{\frac{(\psi+\kappa)(T-t)}{2}}}{(\psi + \kappa)e^{\psi(T-t)} + \psi - \kappa} \right)^{\frac{2\kappa\theta}{\sigma^2}}, \end{aligned}$$

where

$$\psi = \sqrt{\kappa^2 + 2\sigma^2}.$$

## 4.5 CIR++ model

We refer to [BM06] for more details.

The CIR++ model plays the same role for the CIR model as Hull-White's model does for Vasicek's model. It extends the model such that it recovers the market's bond prices. To ensure analytical bond prices it may be specified as follows

$$\begin{aligned} dx(t) &= \kappa(\theta - x(t))dt + \sigma\sqrt{x(t)}dB(t) \\ r(t) &= x(t) + \phi(t) \end{aligned}$$

with  $2\kappa\theta > \sigma^2$ , where

$$\begin{aligned} \phi(t) &= f^M(0, t) - f^{CIR}(0, t) \\ f^{CIR}(0, t) &= \frac{2\kappa\theta(e^{\psi t} - 1)}{2\psi + (\psi + \kappa)(e^{\psi t} - 1)} + x(0) \frac{4\psi^2 e^{\psi t}}{(2\psi + (\psi + \kappa)(e^{\psi t} - 1))^2} \end{aligned}$$

Positivity of interest rates can only be guaranteed when  $\phi(t) > 0$ . This condition is however rather restrictive and in case of negative market forward rates it cannot be fulfilled. In calibrating the model one may choose  $x(0)$  e.g. to stay as close as possible to positive rates. However, this may have a negative impact on fitting to caps and floors.

**Exercise 38.** Show that  $f^{CIR}(0, t)$  equals the current instantaneous forward rate for time  $t$  in the CIR model.

**Exercise 39.** Show that we can set up the Hull-White model in a similar way where

$$\phi(t) = f^M(0, t) - f^{VAS}(0, t)$$

and  $f^{VAS}(0, t)$  denotes the current instantaneous forward rate for time  $t$  in the Vasicek model.

## Chapter 5

# Interest rate market models

We refer to [BM06] for an in-depth discussion on the interest rate market models.

### 5.1 Caplets

Consider two future times  $0 < T_1 < T_2$ . Let  $L(T_1, T_2)$  denote the simply compounded risk free rate fixed at time  $T_1$  over the time interval  $[T_1, T_2]$ . This is often called the LIBOR (London InterBank Offered Rate)rate and represents the average offered interest rate at which banks are willing to trade over the time interval. It means that

$$\frac{1}{P(T_1, T_2)} = 1 + L(T_1, T_2)(T_2 - T_1).$$

This rate is unknown at time  $t < T_1$ . A caplet is an option on this forward rate paying off

$$(T_2 - T_1) \max(0, L(T_1, T_2) - K)$$

at time  $T_2$ .

It can be used to cap the LIBOR rate that has to be paid on a given contract to the strike rate  $K$ .

At time  $t < T_1$  we may consider the forward rate

$$F(t, T_1, T_2) = \frac{1}{T_2 - T_1} \left( \frac{P(t, T_1)}{P(t, T_2)} - 1 \right).$$

This rate converges to the LIBOR rate at time  $T_1$  and hence we can consider the caplet as an option on this forward rate.

Note that the caplet's payoff at time  $T_2$  is equivalent to a payoff at time  $T_1$  of

$$P(T_1, T_2)(T_2 - T_1) \max(0, L(T_1, T_2) - K).$$

This expression is suitable for pricing in the risk neutral world attached to the  $T_2$ -maturity zero coupon bond numeraire. We call the associated measure  $\mathbb{Q}(T_2)$  the  $T_2$ -forward measure. The caplet value at time  $t$  then equals

$$V(t) = P(t, T_2) \mathbb{E}^{\mathbb{Q}(T_2)} [(T_2 - T_1) \max(0, L(T_1, T_2) - K)].$$

Now

$$P(t, T_2)F(t, T_1, T_2) = \frac{P(t, T_1) - P(t, T_2)}{T_2 - T_1}$$

is the price of a tradable asset. So, if we want a market model for the yield curve its price relative to the numeraire must be a martingale. Hence  $F(t, T_1, T_2)$  itself must be a martingale.

The LIBOR forward model is then specified by setting

$$dF(t, T_1, T_2) = \sigma(t)F(t, T_1, T_2)dW^{\mathbb{Q}(T_2)},$$

where  $W^{\mathbb{Q}(T_2)}$  is Brownian motion with respect to  $\mathbb{Q}(T_2)$ .

It follows that the price of the caplet may be computed analytically. This is the justification behind the so-called Black formula (analogue of the Black-Scholes formula for equity calls).

## 5.2 LIBOR market model

The above model specifies the dynamics of one forward rate relative to its forward measure. This is sufficient to price a caplet, or a series of caplets (cap). However, in case we want to specify a model for the complete yield curve we need to consider the dynamics of all forward rates relative to one measure. In the LIBOR market model the yield curve is discretised according to the maturity of the forward rates underlying the caplets, e.g. each 6 months:  $0 = T_0 < T_1 < \dots < T_N$ .

Denote the forward rates by

$$F_k(t) := F(t, T_{k-1}, T_k) = (P(t, T_{k-1})/P(t, T_k) - 1)/\tau_k,$$

where  $\tau_k = T_k - T_{k-1}$ . Then the complete model is specified by the set of SDEs

$$\frac{dF_k(t)}{F_k(t)} = \sigma_k(t)dB^k(t),$$

where each  $B^k(t)$  is Brownian motion with respect to the  $T_k$ -forward measure and  $[B_k, B_l]_t = \rho_{k,l}dt$ .

By the change of numeraire technique we can also express all forward rates with respect to one  $T_i$ -forward measure, yielding

$$\frac{dF_k(t)}{F_k(t)} = \sigma_k(t) \sum_{j=i+1}^k \frac{\rho_{k,j}\sigma_j(t)\tau_j F_j(t)}{1 + \tau_j F_j(t)} dt + \sigma_k(t)dB^k(t) \text{ for } i < k$$

$$\frac{dF_k(t)}{F_k(t)} = -\sigma_k(t) \sum_{j=k+1}^i \frac{\rho_{k,j}\sigma_j(t)\tau_j F_j(t)}{1 + \tau_j F_j(t)} dt + \sigma_k(t) dB^k(t) \text{ for } i > k$$

Note that the correlated Brownian motions for the forward rates can be constructed from a linear combination of just a few independent Brownian motions, called the factors. The initial yield curve is automatically reproduced in the sense that the forward rates according to the grid are respected.

The model can then be used to price any interest rate derivative with payoff times amongst the  $T_k$  via Monte Carlo simulation. We need to choose the  $T_i$ -numeraire such that  $T_i$  is larger than or equal to the final payoff of the derivative. Indeed, the value at time  $t$  of a derivative with only payoff  $V(T)$  at time  $T < T_i$  equals

$$V(t) = P(t, T_i) \mathbb{E}^{\mathbb{Q}(T_i)} \left[ \frac{V(T)}{P(T, T_i)} \right].$$

Note that due to the drift corrections it is necessary to discretise time in order to simulate the SDEs.

### 5.3 Swap Market Model

Consider a swap for which the floating leg pays  $L(T_{j-1}, T_j)\tau_j$  at times  $T_j$  for  $j = \alpha + 1, \dots, \beta$ , where  $L(T_{j-1}, T_j)$  is the LIBOR rate fixed at  $T_{j-1}$ . The fixed leg pays  $K\tau_j$  at times  $T_j$  where  $K$  denotes the swap rate.

At any time  $t < T_\alpha$ , the value of the fixed leg equals

$$\sum_{j=\alpha+1}^{\beta} P(t, T_j) \tau_j K,$$

We may replicate the floating leg payoffs as follows. At time  $t$  we buy a  $T_\alpha$ -maturity zero coupon bond and sell a  $T_\beta$ -maturity zero coupon bond (both with unit face value). At time  $T_\alpha$  we invest the unit face value of the  $T_\alpha$ -bond in  $T_{\alpha+1}$ -zero coupon bonds. This yields the LIBOR rate over the period  $\tau_{\alpha+1}$  at time  $T_{\alpha+1}$  (i.e. the floating leg payoff) plus one unit of currency that we reinvest in a  $T_{\alpha+2}$ -zero coupon bond. We can repeat this investment strategy. At the last period we buy a unit face value  $T_\beta$ -zero coupon bond that cancels our initial short position in this bond. This replicating strategy shows that the value of the floating leg equals

$$P(t, T_\alpha) - P(t, T_\beta).$$

The (forward) swap rate at time  $t$  is the fixed rate that produces equal values for both legs:

$$S_{\alpha,\beta}(t) = \frac{P(t, T_\alpha) - P(t, T_\beta)}{\sum_{j=\alpha+1}^{\beta} P(t, T_j) \tau_j}.$$

Now consider a payer swaption, i.e. the right to enter the swap (as fixed leg payer) at  $T_\alpha$  at swap rate  $K$  (without up front payment). You will only exercise this right when the swap rate  $S_{\alpha,\beta}(T_\alpha) \geq K$ . The value of the swaption then equals the difference in the two fixed legs (at par versus at strike  $K$ ) i.e.

$$\sum_{j=\alpha+1}^{\beta} P(T_\alpha, T_j) \tau_j (S_{\alpha,\beta}(T_\alpha) - K).$$

Hence we can say that the swaption payoff at time  $T_\alpha$  equals

$$\sum_{j=\alpha+1}^{\beta} P(T_\alpha, T_j) \tau_j \max(0, S_{\alpha,\beta}(T_\alpha) - K).$$

It follows that

$$C^{\alpha,\beta}(t) = \sum_{j=\alpha+1}^{\beta} P(t, T_j) \tau_j$$

is a suitable numeraire for pricing the swaption. Note that being a simple portfolio of zero coupon bonds it is a tradable asset. We call this numeraire the swap numeraire and denote its associated measure by  $Q^{\alpha,\beta}$ .

Since by definition the swap rate equals the value of a tradable asset discounted at the swap numeraire, it must be a martingale with respect to its associated risk neutral measure  $Q^{\alpha,\beta}$ .

The swap market model is then specified by setting

$$dS_{\alpha,\beta}(t) = \sigma^{\alpha,\beta} S_{\alpha,\beta}(t) dB^{\alpha,\beta}(t),$$

where  $B^{\alpha,\beta}(t)$  denotes Brownian motion under  $Q^{\alpha,\beta}$ .

The value of the swaption at time  $t$  then equals

$$V(t) = C^{\alpha,\beta}(t) \mathbb{E}^{Q^{\alpha,\beta}} [\max(0, S_{\alpha,\beta}(T_\alpha) - K)].$$

It is clear that due to the model choice a Black-like formula exists for its price.

## 5.4 Multi-curve paradigm in interest rate modeling

In the previous sections we tacitly assumed all interest rate derivatives in an economy can be defined and modeled using one risk free interest rate curve. However, such an approach ignores the credit risk in products like 6M swaps. In this section, we follow [FST10]. They consider fully collateralized trades and distinguish between the risk free rate  $r(t)$  and the collateral rate  $r_c(t)$ . In practice, the SOFR (secured overnight funding rate) is considered a good approximation of the risk free rate for USD (and likewise Ester for EUR, ...). The collateral rate is the rate earned on collateral (i.e. on the marginal account).

For standard market swaps on USD this is nowadays also the SOFR, but it used to be the EFFR (Effective Federal Funds Rate), which is the unsecured overnight rate between banks in USD. For OTC contracts this rate is usually agreed between the counterparties and can be seen as a bank specific unsecured funding rate. For sake of simplicity, we assume it is symmetric.

Assuming perfect continuous collateralization for a claim with value  $h(s)$  at time  $s$ , the collateral account equals  $h(s)$  at time  $s$  and produces investment returns  $(r(s) - r_c(s))h(s)ds$  over  $[s, s + ds]$ . Hence

$$h(t) = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T r(s)ds} h(T) + \int_t^T e^{-\int_t^s r(s)ds} (r(s) - r_c(s))h(s)ds \right]$$

where  $\mathbb{Q}$  denotes the risk neutral measure associated to the risk free bank account  $e^{\int_0^t r(s)ds}$ .

The authors [FST10] show that this implies

$$h(t) = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T r_c(s)ds} h(T) \right]$$

Note that this is the same result as obtained by [Pit10], but here the authors assume the party posting collateral can fund its hedges at a repo rate equal to the risk free rate. As for interest rate hedging one more typically uses Forward Rate Agreements or futures, I guess the collateral rate for posting these would be more suitable and this is usually SOFR, so the distinction disappears when assuming SOFR is risk free. As we've seen from [Pit10] this can be relaxed for general assets by redefining  $\mathbb{Q}$  as a measure in which the asset underlying the claim has a drift equal to the assets repo rate (minus dividend yield).

Let us now consider OIS swaps where the floating leg pays a continuous rate (this is a good approximation of the traded daily compounded rate). Here we assume the rate  $r(t) = r_c(t)$ , i.e. it coincides with the collateral rate. This is for instance the case for SOFR OIS swaps or EFFR OIS swaps. The payoff of a floating leg on a forward period  $[T_{i-1}, T_i]$  equals  $e^{-\int_{T_{i-1}}^{T_i} r(s)ds} - 1$ . So for the OIS swap with fixing dates  $T_i$ ,  $1 \leq i \leq N$  and swap rate  $S_N$  we have

$$S_N(t) \sum_{i=1}^N (T_i - T_{i-1}) \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^{T_i} r(s)ds} \right] = \sum_{i=1}^N \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^{T_i} r(s)ds} \left( e^{-\int_{T_{i-1}}^{T_i} r(s)ds} - 1 \right) \right]$$

or

$$S_N(t) \sum_{i=1}^N (T_i - T_{i-1}) \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^{T_i} r(s)ds} \right] = \sum_{i=1}^N \left( \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^{T_{i-1}} r(s)ds} \right] - \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^{T_i} r(s)ds} \right] \right)$$

Writing  $P(t, T) = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T r(s)ds} \right]$ , we have

$$S_N(t) \sum_{i=1}^N (T_i - T_{i-1}) P(t, T_i) = \sum_{i=1}^N (P(t, T_{i-1}) - P(t, T_i)) = P(t, T_0) - P(t, T_N)$$

For a current swap starting at  $T_0 = 0$  we get

$$S_N(0) \sum_{i=1}^N (T_i - T_{i-1}) P(0, T_i) = 1 - P(0, T_N)$$

One then typically gets the swap rates from the market and derives the discount factors  $P(0, T_N)$  from these quotes. This procedure is called the bootstrap procedure. To do so, one typically interpolates market swap rates such that for each  $N$  (where the  $T_i$  are daily spaced) we get the swap rates  $S_N(0)$ . Then from  $S_1(0)$  derive  $P(0, T_1)$ , next derive  $P(0, T_i)$  from  $S_i(0)$  and  $P(0, T_j)$  for  $i < j$ . By continuous interpolation we can even get the whole OIS curve  $P(0, T)$ . We have seen that the factors  $P(t, T)$  are the relevant discount factors in the pricing formula for fully collateralized trades. That's why we sometimes call the OIS curve the discount curve.

Next we consider swaps on LIBOR rates  $L(T_{i-1}, T_i)$  for tenor  $\Delta \approx T_i - T_{i-1}$ , the rate at which banks are willing to lend money over a time interval of length  $\Delta$ . Note that LIBOR rates are fixed at time  $T_{i-1}$  and paid at time  $T_i$ . The floating leg consists of FRAs (forward rate agreement on a unit notional), i.e. the agreement to receive  $(T_i - T_{i-1})L(T_{i-1}, T_i)$  units of the currency at time  $T_i$ . The value of the (fully collateralized) FRA at time  $t$  then equals

$$\begin{aligned} FRA(t) &= (T_i - T_{i-1}) \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^{T_i} r_c(s) ds} L(T_{i-1}, T_i) \right] \\ &= (T_i - T_{i-1}) P(t, T_i) \mathbb{E}_t^{T_i} [L(T_{i-1}, T_i)]. \end{aligned}$$

Here  $\mathbb{E}_t^{T_i} [\cdot]$  denotes the expectation with respect to the risk neutral measure associated to the  $T_i$ -maturity zero coupon bond. The trick is then to define the LIBOR rate at time  $t$  directly as

$$L_i(t) := \mathbb{E}_t^{T_i} [L(T_{i-1}, T_i)].$$

Like before we have

$$S_N(t) \sum_{i=1}^N (T_i - T_{i-1}) P(t, T_i) = \sum_{i=1}^N P(t, T_i) (T_i - T_{i-1}) L_i(t)$$

Again we can bootstrap the LIBOR rates  $L_i(0)$  from current swap quotes and we can construct the curve  $P_\Delta(t, T)$  such that

$$(T_i - T_{i-1}) L_i(t) = P_\Delta(t, T_{i-1}) - P_\Delta(t, T_i)$$

We call this curve the estimation curve for tenor  $\Delta$ , because the curve gives you the expected (estimated) LIBOR rates.

To price derivatives like caplets on such Libor rates we directly model  $L_i(t)$  in the risk neutral forward measure, e.g. by assuming

$$dL_i(t) = \sigma_i(t)dB_i(t)$$

where  $B_i(t)$  is Brownian motion with respect to the  $T_i$ -forward measure.

Swaptions are treated similar. The swap rate  $S(t)$  is now a weighted sum of these Libor rates

$$S(t) = \frac{\sum_i P(t, T_i)(T_i - T_{i-1})L_i(t)}{\sum_i (T_i - T_{i-1})P(t, T_i)}$$

If we're only interested in swaption pricing we can directly assume the swap rate follows

$$dS(t) = \sigma_S(t)dB(t)$$

where  $B(t)$  is Brownian motion with respect to the swap measure, i.e. the measure associated to the swap numeraire  $\sum_i (T_i - T_{i-1})P(t, T_i)$

# Chapter 6

## Heston's Model

We refer to the original paper of Heston [Hes93] for more details.

### 6.1 Volatility Smile

Black-Scholes' Geometric Brownian Motion assumes a deterministic volatility function  $\sigma(t)$ . This volatility function can be calibrated to observed market prices of call options with a given strike at different maturities. However, according to the model two call options with the same maturity  $T$  but different strikes should have the same Black-Scholes volatility

$$\frac{1}{\sqrt{T}} \sqrt{\int_0^T \sigma^2(t) dt}.$$

However, in practice it turns out that the implied volatility  $\Sigma(T, K)$  also depends on the strike  $K$ . Plotting  $\Sigma(T, K)$  in terms of the strike  $K$  typically produces a smile-like graph (typical for currency options). Depending on the asset the graph may also look like a skew (typical for equity options).

Note that the volatility does not depend on whether the option is a call or put option. This is due to the put-call parity.

The fact that volatility increases for lower strikes can be understood as follows. Suppose first that the implied BS volatility were a flat function of the strike. This means that the lognormal distribution with volatility  $\sigma$  for the asset return coincides with the market's view. However, since the implied volatility is higher for lower strikes, it means that the market assigns relatively more value to the low strike put (compared to the flat volatility). In other words it believes that the (risk neutral) probability of reaching the strike must be higher, i.e. that the left tail of the risk neutral distribution must be fatter.

That is one of the main reasons for looking at models producing fatter tails. One of such models is Heston's model. It achieves this by assuming that the volatility itself is stochastic, a fact that has been observed in historical volatility as well as in BS implied market volatility.

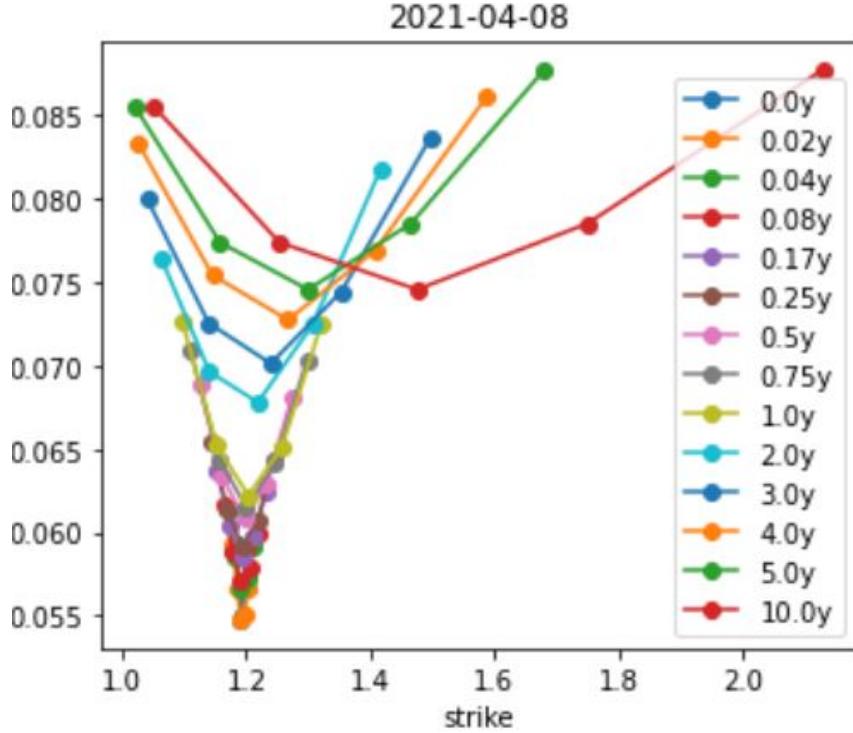


Figure 6.1: Volatility smile curves for USD2EUR at different maturities

## 6.2 Heston's Model

Heston's Model is specified by combining geometric Brownian motion for the asset price with a CIR process for the variance process:

$$\begin{aligned}\frac{dS(t)}{S(t)} &= \mu(t)dt + \sigma(t)dB^S(t) \\ d\sigma^2(t) &= \kappa(\theta - \sigma^2(t))dt + \xi\sigma(t)dB^\sigma(t)\end{aligned}$$

where

$$dB^S(t)dB^\sigma(t) = \rho dt.$$

Note that the volatility  $\sigma^2(0)$  at time 0 needs to be specified. It must be considered as an extra parameter to calibrate. The mean reverting stochastic instantaneous variance  $\sigma^2(t)$  induces a volatility smile on the BS implied volatility curve in terms of the strike (when  $\rho = 0$ ). In equity markets one often sets  $\rho < 0$  in order to arrive at a volatility skew.

The Heston model has the following interesting features:

- it has volatility clustering: the variances of logarithmic returns of the asset price over consecutive time intervals are positively correlated
- it explains the volatility smile/skew
- it is relatively easy to simulate (by discretisation)
- it has quasi-analytic formulae for plain vanilla call and put options

For a risk neutral Heston model  $\mu(t)$  must be the riskfree rate. We are then left with 5 parameters to calibrate:  $\kappa$ ,  $\theta$ ,  $\xi$ ,  $\rho$  and the non-observable  $\sigma(0)$ . It is quite challenging to calibrate the model to plain vanilla call (or put) options as you need to minimize differences between modeled and observed prices in a five dimensional space. One typically minimises the sum of squares of these differences. Optimisation algorithms usually only find a local minimum of this sum function. So it may be necessary to combine them with a grid search. Moreover there may be many solutions yielding approximately the same sums with quite different parameters. In particular, the calibration may become unstable for small changes in the input call prices. Another approach first calibrates a subset of parameters before calibrating the other parameters. You may for instance use variance swaps to calibrate  $\kappa$ ,  $\theta$  and  $\sigma(0)$ . The call prices are then used to further calibrate  $\rho$  and  $\xi$ . This approach may yield more stable parameters. In general it is known that the Heston model can fit the volatility smile for longer maturity options quite well. Short maturity options are harder to calibrate, since the model does not allow for jumps.

**Exercise 40.** Show that

$$E \left( \frac{1}{T} \int_0^T \sigma^2(t) dt \right) = \theta + (\sigma^2(0) - \theta) \frac{1 - e^{-\kappa T}}{\kappa T}.$$

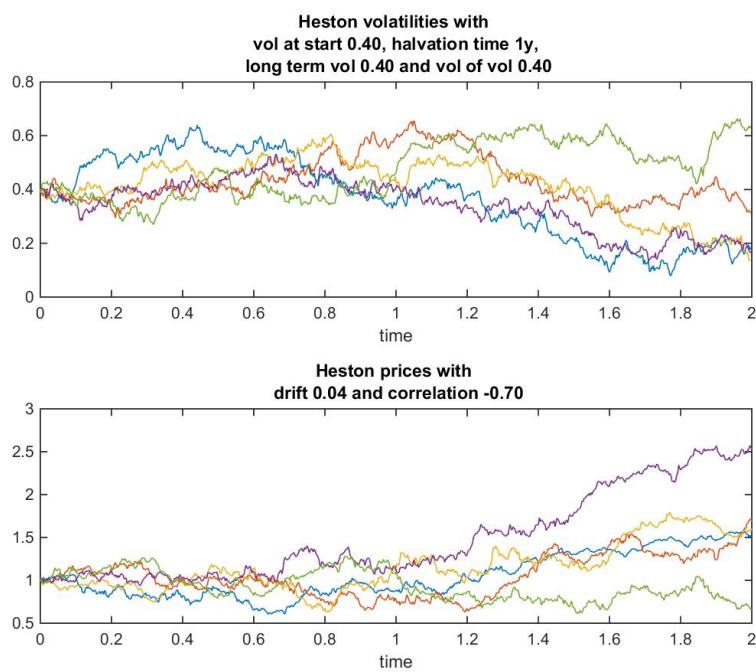


Figure 6.2: 5 sample paths of a Heston process

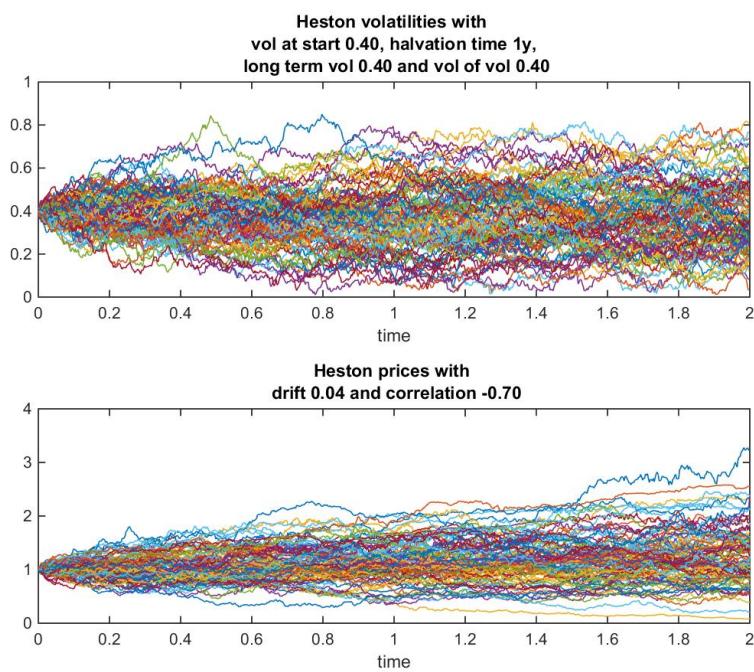


Figure 6.3: 100 sample paths of a Heston process

# Chapter 7

## Lévy Processes

A nice introduction to Lévy processes can be found in [Sch03]. More extensive excellent references are [Sat13] and [CT04].

### 7.1 Poisson Process

A Poisson process  $N_\lambda(t)$  on a probability space  $(\Omega, \mathcal{F}, P)$  is a cadlag stochastic process such that

- $N_\lambda(0) = 0$
- its increments over disjoint time intervals are independent
- and follow a Poisson distribution with parameter  $\lambda t$  where  $t$  denotes the length of the time interval.

Hence  $E(N_\lambda(t)) = \lambda t$  and we call  $\lambda$  the expectation rate.

Hence

$$P(N_\lambda(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

Note that the characteristic function  $\Phi_{N_\lambda(1)}(\theta)$  of  $N_\lambda(1)$  equals

$$\Phi_{N_\lambda(1)}(\theta) := E(e^{i\theta N_\lambda(1)}) = e^{\lambda(e^{i\theta}-1)}.$$

and moreover

$$\Phi_{N_\lambda(t)}(\theta) = (\Phi_{N_\lambda(1)}(\theta))^t.$$

**Exercise 41.** Let  $\tau$  denote the first jump time of the Poisson process  $N_\lambda(t)$ . Show that  $\tau \sim \text{Exp}(\lambda)$ .

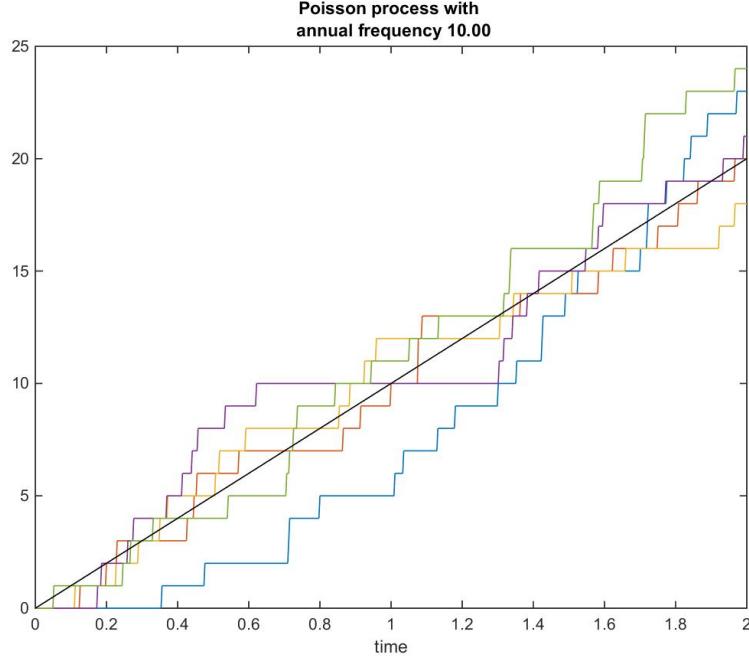


Figure 7.1: 5 sample paths of a Poisson process

**Exercise 42.** Let  $\tau_n, n = 1, 2, \dots$  be a sequence of i.i.d  $\text{Exp}(\lambda)$  distributed random variables. Define a process  $N_\lambda(t)$  by setting

$$N_\lambda(t) = N \Leftrightarrow \sum_{n=1}^N \tau_n \leq t \text{ and } \sum_{n=1}^{N+1} \tau_n > t.$$

Show that  $N_\lambda(t)$  is a Poisson process with expectation rate  $\lambda$ . Hint: prove by induction on  $N$  that for all  $N \geq 0$

$$P\left(\sum_{n=1}^{N+1} \tau_n > t\right) = e^{-\lambda t} \left( \sum_{n=0}^N \frac{(\lambda t)^n}{n!} \right).$$

## 7.2 Compound Poisson process

A compound Poisson process  $X(t)$  on a probability space  $(\Omega, \mathcal{F}, P)$  is like a Poisson process with jumps of different sizes according to a jump size probability measure  $J$ . Thus

$$X(t) = \sum_{n=1}^{N_\lambda(t)} J_n,$$

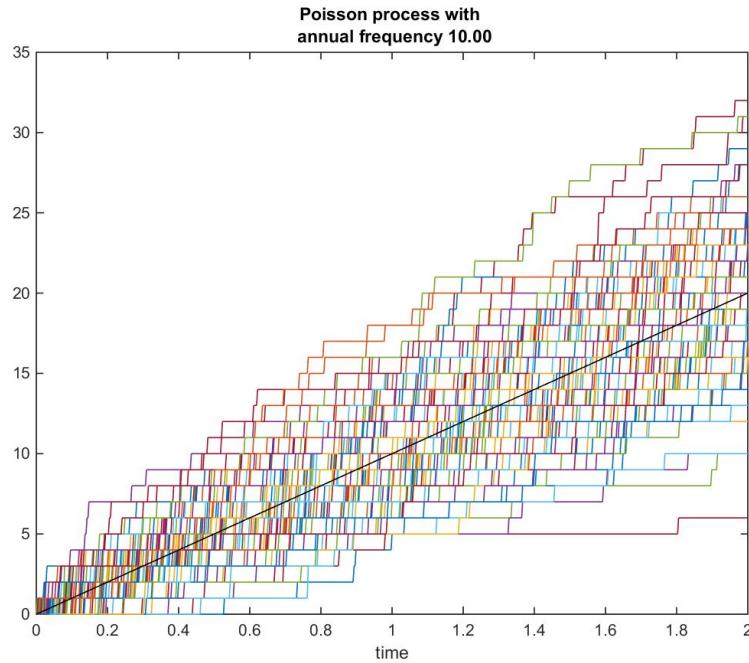


Figure 7.2: 100 sample paths of a Poisson process

where  $N_\lambda(t)$  is a Poisson process with parameter  $\lambda$  and the  $J_n$  are independent copies of a random variable with  $P(J_n \in A) = J(A)$ . Here  $J$  is a jump measure and must satisfy  $J(\{0\}) = 0$ .

Again this process has independent increments over disjoint time intervals and the distribution of increments over time intervals of the same length is identical.

Denote the characteristic function of  $J$  by  $\Phi_J(\theta)$ . The characteristic function of  $X(1)$  may then be computed as follows

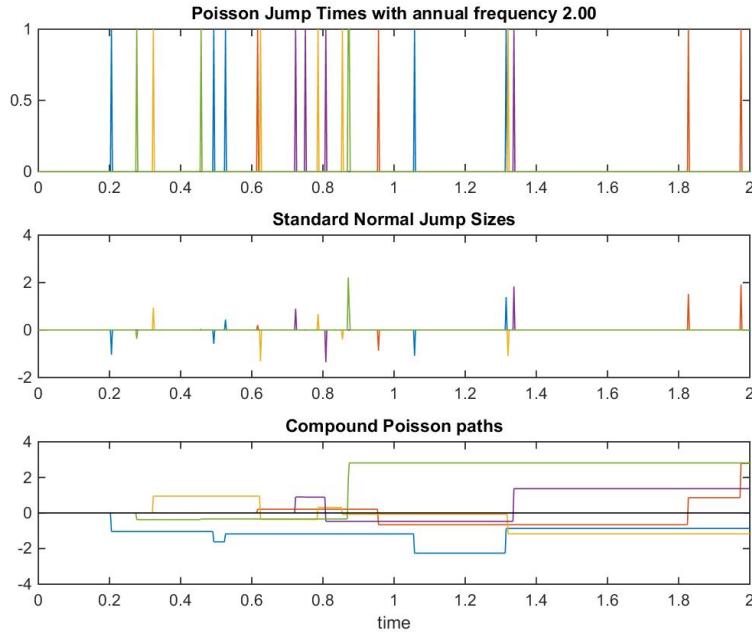


Figure 7.3: 5 sample paths of a compound Poisson process

$$\begin{aligned}
 \Phi_{X(1)}(\theta) &= E(e^{i\theta X(1)}) \\
 &= \sum_{n=1}^{\infty} P(N(1) = n) \Phi_{\sum_{i=1}^n J_i}(\theta) \\
 &= \sum_{n=1}^{\infty} P(N(1) = n) (\Phi_J(\theta))^n \\
 &= \sum_{n=1}^{\infty} \frac{(\lambda \Phi_J(\theta))^n}{n!} e^{-\lambda} \\
 &= e^{\lambda(\Phi_J(\theta)-1)} \\
 &= \Phi_{N_\lambda(1)}(-i \ln \Phi_J(\theta))
 \end{aligned}$$

Let  $A$  be a measurable subset of  $\mathbb{R} \setminus \{0\}$ , and let  $N_A(t)$  denote the process counting the number of jumps of  $X(t)$  of size in  $A$ . Then, a similar computation, yields

$$\begin{aligned}
\Phi_{N_A(1)}(\theta) &= E(e^{i\theta N_A(1)}) \\
&= \sum_{n=0}^{\infty} P(N_\lambda(1) = n) E(e^{i\theta \sum_{j=1}^n 1_{\{J_j \in A\}}} | N_\lambda(1) = n) \\
&= \sum_{n=0}^{\infty} P(N_\lambda(1) = n) (1 - J(A) + J(A)e^{i\theta})^n \\
&= e^{\lambda J(A)(e^{i\theta} - 1)}
\end{aligned}$$

Thus  $N_A(t)$  is a Poisson process with expectation  $\lambda J(A)t$ .

**Exercise 43.** Show that a compound Poisson process reduces to a Poisson process if and only if the jump measure is the Dirac measure at 1, i.e.  $J(\{1\}) = 1$ .

### 7.3 Lévy Process

We have seen that both Brownian motion with drift and compound Poisson processes are processes with independent and identically distributed increments over disjoint time intervals of the same length. This is essentially the definition of a Lévy process and Lévy showed that each such process is essentially the sum of Brownian motion with drift and a (slight generalisation) of a compound Poisson process.

**Lévy process:** A cadlag stochastic process  $X(t)$  on a probability space  $(\Omega, \mathcal{F}, P)$  is a Lévy process if

- $X(0) = 0$  almost surely
- it has independent increments over disjoint time intervals
- it has identical increments over time intervals of the same length
- it is continuous in probability: for all  $\epsilon > 0$  and  $t \geq 0$

$$\lim_{dt \rightarrow 0} P(|X(t+dt) - X(t)| > \epsilon) = 0.$$

**Lévy-Khintchine representation:** The characteristic function  $\Phi_{X(1)}(\theta)$  of a Lévy process  $X(t)$  is given by

$$\exp \left( ai\theta - \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R} \setminus \{0\}} (e^{i\theta x} - 1 - i\theta x 1_{\{|x|<1\}}(x)) \Pi(dx) \right),$$

where  $a, \sigma \in \mathbb{R}$  and  $\Pi$  is a measure on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  satisfying

$$\int_{\mathbb{R} \setminus \{0\}} \min(1, x^2) \Pi(dx) < \infty.$$

The characteristic function is completely determined by the so-called Lévy triplet  $(a, \sigma, \Pi)$ .

Note the indicator function  $1_{\{|x|<1\}}(x)$ . It allows the density of the jump measure to increase to  $\infty$  around  $x = 0$ . One could as well use the indicator function  $1_{|x|<b}(x)$  for any  $b > 0$ . This only changes the value of  $a$ . Note that the integrand is of order  $x^2$  around  $x = 0$ . Away from 0 the integrand is bounded. Hence, by the extra condition, the integral exists.

In case of a compound Poisson process with frequency parameter  $\lambda$  and jump size measure  $J$ , we can define  $\Pi = \lambda J$ . We then get:

$$\begin{aligned} \ln \Phi_{X(1)}(\theta) &= \lambda(\Phi_J(\theta) - 1) \\ &= \int_{\mathbb{R} \setminus \{0\}} (e^{i\theta x} - 1) \lambda J(dx) \\ &= i\theta \int_{\mathbb{R} \setminus \{0\}} x 1_{\{|x|<1\}}(x) \Pi(dx) \\ &\quad + \int_{\mathbb{R} \setminus \{0\}} (e^{i\theta x} - 1 - i\theta x 1_{\{|x|<1\}}(x)) \Pi(dx) \\ &= i\theta a + \int_{\mathbb{R} \setminus \{0\}} (e^{i\theta x} - 1 - i\theta x 1_{\{|x|<1\}}(x)) \Pi(dx) \end{aligned}$$

where

$$a = \int_{\mathbb{R} \setminus \{0\}} x 1_{\{|x|<1\}}(x) \lambda J(dx).$$

The Lévy triplet thus equals  $(a, 0, \lambda J)$ .

It is clear from the representation theorem that each Lévy process can be decomposed into drift plus Brownian motion plus a generalised compound Poisson process.

Similarly as before, for every measurable set  $A$  with  $\Pi(A) < \infty$  the process  $X_A(t)$  of jumps of size in  $A$  is a Poisson process with mean  $\Pi(A)t$ . Hence the measure  $\Pi$  is called the jump intensity measure.

**Infinite divisibility:** It follows from the definition that the distribution of a Lévy process at any time  $t$  is infinitely divisible. It means that for arbitrary  $n$  it can be obtained as the sum of  $n$  independent identically distributed random variables. The converse also holds: given an infinitely divisible distribution, then there exists a Lévy process  $X(t)$  for which  $X(1)$  has this distribution. It can be used to construct Lévy processes.

**Exercise 44.** Show that when  $\Pi(\mathbb{R} \setminus \{0\}) = \infty$ , the process has almost surely infinitely many jumps over any time interval of positive length.

**Exercise 45.** Show that when

$$\int_{\mathbb{R} \setminus \{0\}} \min(1, x) \Pi(dx) < \infty,$$

we have

$$\Phi_{X(1)}(\theta) = \exp \left( ci\theta - \frac{1}{2} \sigma^2 \theta^2 + \int_{\mathbb{R} \setminus \{0\}} (e^{i\theta x} - 1) \Pi(dx) \right),$$

for some constant  $c$ .

## 7.4 Gamma process

Consider a random variable  $X$  with gamma distribution with mean  $\mu$  and variance  $\nu$ . Thus

$$P(X \leq x) = \frac{1}{\Gamma(\mu^2/\nu)(\nu/\mu)^{\mu^2/\nu}} \int_0^x e^{-u\mu/\nu} u^{\mu^2/\nu-1} du.$$

and its characteristic function equals

$$\Phi_X(\psi) = E(e^{i\psi X}) = (1 - i\psi\nu/\mu)^{-\mu^2/\nu}.$$

This distribution is infinitely divisible since for every  $N$  we have

$$(1 - i\psi\nu/\mu)^{-\mu^2/\nu} = \left( (1 - i\psi(\nu/N)/(\mu/N))^{-(\mu/N)^2/(\nu/N)} \right)^N.$$

Hence we can construct the gamma process  $\Gamma_\nu(t)$  as a Lévy process such that its distribution at time 1 is a gamma distribution with mean 1 and variance  $\nu$ . It follows that the distribution of  $\Gamma_\nu(t)$  is a gamma distribution with mean  $t$  and variance  $\nu t$ .

One may check that the Lévy intensity measure equals

$$\Pi(dx) = \frac{1}{\nu} x^{-1} e^{-x/\nu} dx.$$

Indeed,

$$\begin{aligned} \int_0^\infty \frac{1}{\nu} x^{-1} (e^{i\psi x} - 1) e^{-x/\nu} dx &= \frac{1}{\nu} \int_0^\infty x^{-1} \sum_{n=1}^\infty \frac{(i\psi x)^n}{n!} e^{-x/\nu} dx \\ &= \frac{1}{\nu} \sum_{n=1}^\infty \frac{(i\psi\nu)^n}{n} \\ &= -\frac{1}{\nu} \ln(1 - i\psi\nu). \end{aligned}$$

Note that the Gamma process is not a straight compound Poisson process, since for every  $\epsilon > 0$  it has infinitely many jumps of size smaller than  $\epsilon$  in any time interval of positive length.

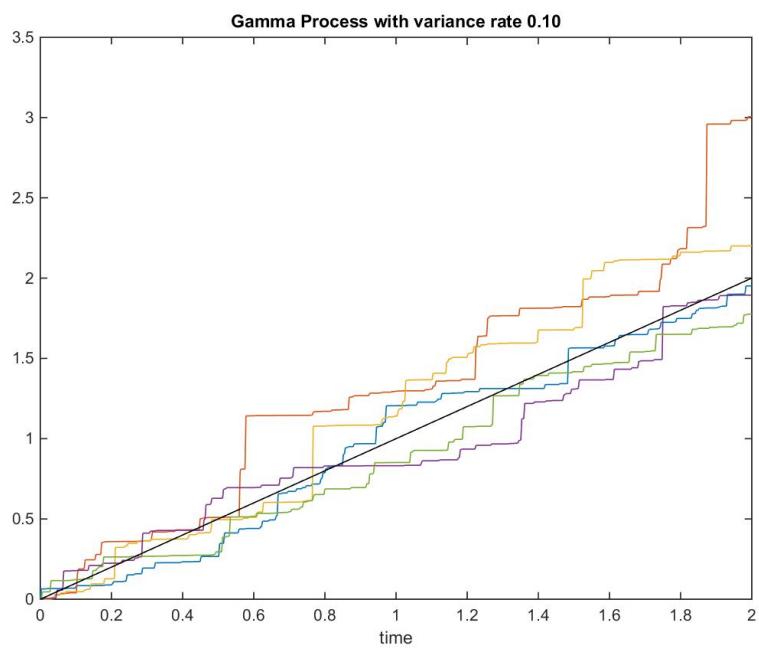


Figure 7.4: 5 sample paths of a Gamma process

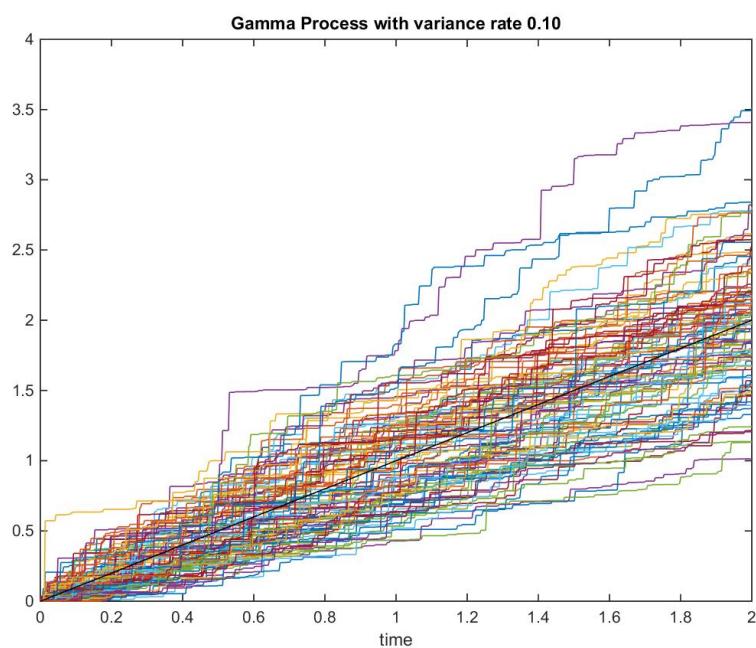


Figure 7.5: 100 sample paths of a Gamma process

Jump intensity density of Gamma process with nu=1

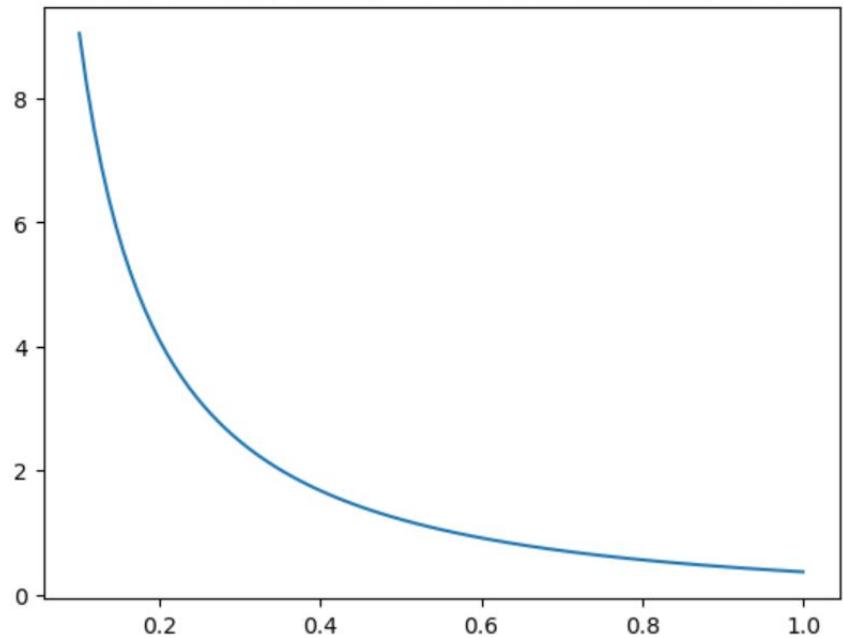


Figure 7.6:

## 7.5 Variance Gamma process

We refer to [CCM98] for more details about the variance gamma process.

The variance gamma process is obtained by replacing time  $t$  in a Brownian motion with drift by  $\Gamma_\nu(t)$  :

$$X(t) = \theta\Gamma_\nu(t) + \sigma B(\Gamma_\nu(t)),$$

where  $\Gamma_\nu(t)$  and  $B(t)$  are independent.

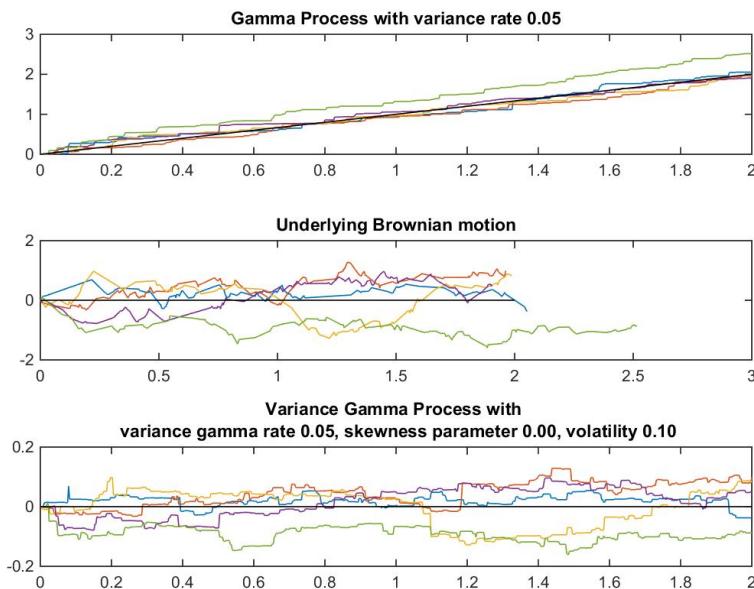


Figure 7.7: 5 sample paths of a variance gamma process

Since Brownian motion with drift and the Gamma process are independent Lévy processes, the variance gamma process is a Lévy process as well. Moreover, it is easy to compute its characteristic function:

$$E(e^{i\psi X(t)}) = \left(1 - i(\psi\theta + i\frac{\sigma^2\psi^2}{2})\nu\right)^{-\frac{1}{\nu}}.$$

Note that the quadratic variation of the Brownian motion  $B(\Gamma_\nu(t))$  is the gamma process  $\Gamma_\nu(t)$ . Hence the name of the process.

When  $\theta = 0$  the process is symmetric. In finance one often uses the variance gamma process to model logarithmic returns of equity prices. In such cases one may take  $\theta < 0$  in order to get fatter left side tails, or to model the volatility skew.

## Chapter 8

# Time Changed Lévy Processes

### 8.1 Chronometers

The increments of Lévy processes over disjoint time intervals of the same length are i.i.d. However, in stock markets we observe volatility clustering, meaning that in turmoil periods the distribution of logarithmic returns has larger variance than in calm periods. In more mathematical terms, the distribution of logarithmic returns is conditional on the available information. We can construct stochastic processes exhibiting volatility clustering by replacing the time  $t$  in a given stochastic process  $X(t)$  by an independent increasing process  $\tau(t)$  starting at 0. The result is a time-changed process

$$Y(t) = X(\tau(t)).$$

We call the process  $\tau(t)$  a chronometer. A simple example is a gamma process. However, by applying an independent gamma process as chronometer to a Lévy process we do not obtain a process with volatility clustering. Indeed, the resulting process is again a Lévy process.

More interesting chronometers may be obtained by integrating a mean reverting process such as Vasicek's process:

$$\tau(t) = \int_0^t x(s)ds$$

where

$$dx(s) = \kappa(\theta - x(s)) + \sigma dB(s).$$

Here  $x(s)$  can be interpreted as the speed of time.

An interesting case arises when  $X$  is a Lévy process and the characteristic function of the chronometer is also known:

$$\Phi_X(u, t) = \mathbb{E} \left[ e^{iuX(t)} \right]$$

$$\Phi_\tau(u, t) = \mathbb{E} \left[ e^{iu\tau(t)} \right]$$

We then get

$$\begin{aligned}\Phi_Y(u, t) &= \mathbb{E} \left[ e^{iuY(t)} \right] \\ &= \mathbb{E} \left[ e^{iuX(\tau(t))} \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ e^{iuX(\tau(t))} | \tau(t) \right] \right] \\ &= \mathbb{E} [\Phi_X(u, \tau(t))] \\ &= \mathbb{E} \left[ e^{\tau(t) \log \Phi_X(u, 1)} \right] \\ &= \Phi_\tau(-i \log \Phi_X(u, 1), t)\end{aligned}$$

In the next section we explain how to use this characteristic function to compute plain vanilla call prices.

## 8.2 Pricing plain vanilla calls via characteristic functions

In this section we follow [CM99] quite closely.

Our goal is to price a plain vanilla call option on a non-dividend paying stock given the characteristic function of its risk neutral density at maturity (with respect to the bank account numeraire). Let  $K = e^k$  denote the strike and  $T$  the maturity. Denote the density at maturity in terms of  $s = \log(S(T))$  by  $q(s)$ , where  $S(T)$  denotes the stock price at maturity. For sake of simplicity we assume a constant risk-free rate  $r$ .

The value of the call option then equals

$$C(k) = \mathbb{E} \left[ e^{-r(T)} \max(0, e^s - e^k) \right] = \int_k^\infty e^{-rT} (e^s - e^k) q(s) ds.$$

We would like to take the Fourier transform

$$\int_{-\infty}^{\infty} e^{ivk} C(k) dk,$$

but to this end  $C(k)$  needs to be square integrable. However, as  $C(k)$  converges to  $S(0)$  as  $k \rightarrow -\infty$ , it is not square integrable. For this reason we adapt  $C(k)$  by setting

$$c(k) = C(k) e^{\alpha k}.$$

We will show in a moment that the Fourier transform

$$\Psi(v) = \int_{-\infty}^{\infty} e^{ivk} c(k) dk$$

exists for some suitable  $\alpha > 0$ . We have

$$\begin{aligned}\Psi(v) &= \int_{-\infty}^{\infty} e^{ivk} c(k) dk \\ &= \int_{-\infty}^{\infty} e^{ivk} e^{\alpha k} \int_k^{\infty} e^{-rT} (e^s - e^k) q(s) ds dk \\ &= \int_{-\infty}^{\infty} e^{-rT} q(s) \left( \int_{-\infty}^s e^s e^{(iv+\alpha)k} dk - \int_{-\infty}^s e^{(iv+\alpha+1)k} dk \right) ds \\ &= e^{-rT} \int_{-\infty}^{\infty} q(s) \left( \frac{e^{(iv+\alpha+1)s}}{iv + \alpha} - \frac{e^{(iv+\alpha+1)s}}{iv + \alpha + 1} \right) ds \\ &= e^{-rT} \left( \frac{\Phi(v - (\alpha + 1)i)}{iv + \alpha} - \frac{\Phi(v - (\alpha + 1)i)}{iv + \alpha + 1} \right) \\ &= e^{-rT} \frac{\Phi(v - (\alpha + 1)i)}{(iv + \alpha)(iv + \alpha + 1)}\end{aligned}$$

where  $\Phi(v) = \mathbb{E}[e^{ivs}]$  denotes the characteristic function of the risk neutral density of  $s$ . Note that

$$|\Psi(v)| \leq |\Psi(0)| = e^{-rT} \frac{\Phi(-(\alpha + 1)i)}{\alpha(\alpha + 1)},$$

so the Fourier transform exists if

$$|\Phi(-(\alpha + 1)i)| = |\mathbb{E}[S(T)^{\alpha+1}]| < \infty.$$

Applying the inverse Fourier transform, we get

$$\begin{aligned}
C(k) &= e^{-\alpha k} c(k) \\
&= e^{-\alpha k} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ivk} \Psi(v) dv \\
&= e^{-\alpha k} \frac{e^{-rT}}{2\pi} \int_{-\infty}^{\infty} e^{-ivk} \frac{\Phi(v - (\alpha + 1)i)}{(iv + \alpha)(iv + \alpha + 1)} dv \\
&= \frac{e^{k-rT}}{2\pi i} \int_{1+\alpha-i\infty}^{1+\alpha+i\infty} e^{-ku} \frac{\Phi(-iu)}{(u-1)u} du \\
&= \frac{e^{k-rT}}{2\pi i} \left( \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} e^{-ku} \frac{\Phi(-iu)}{(u-1)u} du + 2\pi i e^{-k} \Phi(-i) \right) \\
&= -Ke^{-rT} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ku} \frac{\Phi(-iu)}{\frac{1}{4}+y^2} dy + S(0) \\
&= S(0) - Ke^{-rT} \int_0^1 e^{-ku} \Phi(-iu) dx
\end{aligned}$$

where we have used the following substitutions

$$\begin{aligned}
u &= \alpha + 1 + iv \\
u &= \frac{1}{2} + iy \\
y &= \frac{1}{2} \tan\left(\frac{\pi}{2}x\right)
\end{aligned}$$

and we got rid of  $\alpha$  by applying the residue theorem from complex analysis. Note that the integral in the obtained formula can easily be computed by numerical methods once the characteristic function is known in analytical form.

**Exercise 46.** Show that

$$\int_{1+\alpha-i\infty}^{1+\alpha+i\infty} \frac{f(u)}{(u-1)u} du = \int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \frac{f(u)}{(u-1)u} du + 2\pi i f(1)$$

where  $f(u)$  is an analytic function on the domain  $\{u | Re(u) \in [\frac{1}{2}, 1 + \alpha]\}$ . Hint: use the residue theorem and show that  $\int_{\frac{1}{2}\pm iR}^{1+\alpha\pm iR} \frac{f(u)}{(u-1)u} du \rightarrow 0$  for  $|R| \rightarrow \infty$ .

# Chapter 9

## Appendices

### 9.1 Density function of the CIR short rate

Let the CIR process be given by the SDE

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{r(t)}dB(t).$$

We will now derive an analytic formula for its density based on a paper of Feller [Fel51].

Setting  $a = \sigma^2/2$ ,  $b = -\kappa$ ,  $c = \kappa\theta$  and  $x = r(t)$  the PDE for the density  $\rho(r, t)$  of  $r(t)$  at time  $t$  can be written as

$$\rho_t = -((bx + c)\rho)_x + (ax\rho)_{xx}$$

where we use subscript notation for partial derivatives.

To solve the PDE take the Laplace transform

$$\omega(s, t) = \mathcal{L}(\rho) = \int_0^\infty e^{-sx} \rho(x, t) dx.$$

The PDE then transforms to

$$\begin{aligned} \mathcal{L}(\rho_t) &= \omega_t \\ &= \mathcal{L}((-bx - c)\rho + (ax\rho)_x) + \lim_{x \rightarrow 0} ((bx + c)\rho - (ax\rho)_x) \\ &= s(as\mathcal{L}(x\rho) - b\mathcal{L}(x\rho) - c\mathcal{L}(\rho)) + f(t) \\ &= s(-as + b)\omega_s - sc\omega + f(t) \end{aligned}$$

We thus need to solve

$$dt = \frac{ds}{s(as - b)} = \frac{d\omega}{f(t) - cs\omega}$$

The first equation can be solved by writing

$$\frac{1}{s(as-b)} = \frac{a}{b(as-b)} - \frac{1}{bs}$$

and integrating out  $s$ . We get

$$e^{-bt} = C_1 \frac{s}{as-b}.$$

Substituting this in the second equation we need to solve

$$\omega_t = f(t) - sc\omega = f(t) - c \frac{b}{a - C_1 e^{bt}} \omega.$$

Write

$$\lambda = \omega g,$$

then

$$\frac{\lambda_t}{g} = \omega_t + \omega \frac{g_t}{g}$$

So by solving

$$\frac{g_t}{g} = \frac{cb}{a - C_1 e^{bt}}$$

we get

$$g = D |C_1 - ae^{-bt}|^{-c/a},$$

for some constant  $D$  and

$$\lambda_t = f(t)g,$$

so

$$\lambda = D \left( \int_0^t f(\tau) |C_1 - ae^{-b\tau}|^{-c/a} d\tau + C_2 \right)$$

for some constant  $C_2$  and hence

$$\omega = |C_1 - ae^{-bt}|^{c/a} \left( \int_0^t f(\tau) |C_1 - ae^{-b\tau}|^{-c/a} d\tau + C_2 \right).$$

We now look for solutions  $\rho(x, t)$  with initial condition equal to the Dirac measure in  $r(0)$

$$\rho(x, 0) = \delta_{r(0)}(x)$$

whose Laplace transform equals

$$\omega(s, 0) = \pi(s) = e^{-r(0)s}.$$

Hence

$$\pi(s) = |C_1 - a|^{c/a} C_2$$

from which

$$C_2 = \pi(s)|C_1 - a|^{-c/a} = \pi\left(\frac{b}{a - C_1}\right)|C_1 - a|^{-c/a}.$$

We thus get

$$\begin{aligned} \omega(s, t) &= |e^{-bt}C_1 - ae^{-bt}|^{c/a} \left( \int_0^t f(\tau)|C_1 - ae^{-b\tau}|^{-c/a} d\tau + C_2 \right) \\ &= |e^{-bt}C_1 - ae^{-bt}|^{c/a} \left( \int_0^t f(\tau)|C_1 - ae^{-b\tau}|^{-c/a} d\tau + \pi\left(\frac{b}{a - C_1}\right)|C_1 - a|^{-c/a} \right) \\ &= |e^{-bt}\frac{b}{s}|^{c/a} \left( \pi\left(\frac{b}{a(1 - e^{-bt}) + e^{-bt}\frac{b}{s}}\right)|a(1 - e^{-bt}) + e^{-bt}\frac{b}{s}|^{-c/a} \right. \\ &\quad \left. + \int_0^t f(\tau)|a(e^{-bt} - e^{-b\tau}) - e^{-bt}\frac{b}{s}|^{-c/a} d\tau \right) \\ &= \left( \frac{b}{as(e^{bt} - 1) + b} \right)^{c/a} \pi\left(\frac{bse^{bt}}{as(e^{bt} - 1) + b}\right) + \int_0^t f(\tau)(as(e^{b(t-\tau)} - 1) + b)^{-c/a} d\tau \end{aligned}$$

We now look for a solution with  $f(t) = 0$ . This can be obtained by taking the inverse Laplace transform.

$$\begin{aligned} u(r(t)) &= \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{sr(t)} \omega(s, t) ds \\ &= \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} e^{sr(t)} \left( \frac{b}{\gamma s + b} \right)^{c/a} \exp\left(-r(0)\frac{bse^{bt}}{\gamma s + b}\right) ds \\ &= \frac{1}{2\pi i} e^{-\frac{b}{\gamma}(r(t) + r(0)e^{bt})} \left( \frac{b}{\gamma} \right)^{c/a} \int_{b/\gamma - i\infty}^{b/\gamma + i\infty} e^{zr(t)} z^{-c/a} e^{\alpha z^{-1}} dz \\ &= \frac{1}{2\pi i} e^{-\frac{b}{\gamma}(r(t) + r(0)e^{bt})} \left( \frac{b}{\gamma} \right)^{c/a} \left( \frac{\alpha}{r(t)} \right)^{-\frac{c}{2a} + \frac{1}{2}} \int_{b/\gamma \sqrt{\frac{r(t)}{\alpha}} - i\infty}^{b/\gamma \sqrt{\frac{r(t)}{\alpha}} + i\infty} e^{\sqrt{\alpha r(t)}(w + \frac{1}{w})} w^{-c/a} dw \\ &= \frac{1}{2\pi i} e^{-\frac{b}{\gamma}(r(t) + r(0)e^{bt})} \left( \frac{b}{\gamma} \right)^{c/a} \left( \frac{\alpha}{r(t)} \right)^{-\frac{c}{2a} + \frac{1}{2}} I_q(2\sqrt{\alpha r(t)}) \\ &= de^{-u-v} \left( \frac{v}{u} \right)^{\frac{a}{2}} I_q(2\sqrt{uv}) \end{aligned}$$

where we used the following substitutions

$$\begin{aligned}
\gamma &= a(e^{bt} - 1) \\
z &= s + \frac{b}{\gamma} \\
\alpha &= r(0)e^{bt}\left(\frac{b}{\gamma}\right)^2 \\
w &= \sqrt{\frac{r(t)}{\alpha}}z \\
a &= \sigma^2/2 \\
b &= -\kappa \\
c &= \kappa\theta \\
d &= \frac{2\kappa}{\sigma^2(1 - e^{-\kappa(t)})} \\
u &= dr(0)e^{-\kappa(t)} \\
v &= dr(t) \\
q &= \frac{2\kappa\theta}{\sigma^2} - 1
\end{aligned}$$

and  $I_q$  denotes the modified Bessel function of the first kind of order  $q$ .

## 9.2 Solving the Feynman-Kac equation for zero coupon bond prices in a CIR model

In the CIR model the prices of zero coupon bonds  $P(t, T)$  can be obtained by solving the Feynman-Kac equation for the bond price. As in Vasicek's model we assume an affine bond price structure , i.e.

$$P(t, T) = A(t, T)e^{-B(t, T)r(t)}$$

for smooth functions  $A(t, T)$  and  $B(t, T)$ . This yields the following system of ordinary PDEs

$$\begin{aligned}
1 &= -\frac{\partial B(t, T)}{\partial t} + B(t, T)\kappa + \frac{1}{2}\sigma^2 B(t, T)^2 \\
0 &= \frac{\partial A(t, T)/\partial t}{A(t, T)} - B(t, T)\kappa\theta
\end{aligned}$$

The first equation is a so-called Riccati equation and can be solved by substituting

$$B = \frac{2}{\sigma^2}v$$

and

$$v = -\frac{u_t}{u},$$

where we used subscript notation for partial derivatives.  
This yields

$$u = C_1 e^{\gamma_+ t} + C_2 e^{\gamma_- t}$$

for some constants  $C_1$  and  $C_2$  where

$$\gamma_{\pm} = \frac{\kappa \pm \psi}{2}$$

and

$$\psi = \sqrt{\kappa^2 + 2\sigma^2}.$$

It follows that

$$B(t, T) = \frac{-e^{-\psi t} \frac{\kappa - \psi}{2} - C \frac{\kappa + \psi}{2}}{\frac{\sigma^2}{2} (e^{-\psi t} + C)}$$

where  $C$  must be solved by requiring  $B(T, T) = 0$ . Thus

$$C = \frac{\psi - \kappa}{\psi + \kappa} e^{-\psi T}$$

and

$$B(t, T) = \frac{2(e^{\psi(T-t)} - 1)}{(\psi + \kappa)e^{\psi(T-t)} + \psi - \kappa}.$$

To solve  $A(t, T)$  note that the second equation can be rewritten as

$$\frac{\partial \ln A(t, T)}{\partial t} = \kappa \theta B(t, T),$$

and hence integrating out we get

$$\begin{aligned} \ln A(t, T) &= \kappa \theta \int_T^t \frac{2(e^{\psi(T-s)} - 1)}{(\psi + \kappa)e^{\psi(T-s)} + \psi - \kappa} ds + \ln C \\ &= -\frac{2\kappa\theta}{\psi} \int_1^{e^{\psi(T-t)}} \frac{u-1}{((\psi + \kappa)u + \psi - \kappa)u} du + \ln C \\ &= -\frac{2\kappa\theta}{\psi(\psi - \kappa)} \int_1^{e^{\psi(T-t)}} \left( \frac{1}{u} - \frac{2\psi}{(\psi + \kappa)u + \psi - \kappa} \right) du + \ln C \end{aligned}$$

Hence

$$A(t, T) = D \left( \frac{e^{\psi(T-t)}}{((\psi + \kappa)e^{\psi(T-t)} + \psi - \kappa)^{\frac{2\psi}{\psi + \kappa}}} \right)^{-\frac{2\kappa\theta}{\psi(\psi + \kappa)}},$$

where the constant  $D$  is determined by requiring  $A(T, T) = 1$ . It follows that

$$D = (2\psi)^{\frac{2\kappa\theta}{\sigma^2}}$$

and

$$A(t, T) = \left( \frac{2\psi e^{\frac{(\psi + \kappa)(T-t)}{2}}}{(\psi + \kappa)e^{\psi(T-t)} + \psi - \kappa} \right)^{\frac{2\kappa\theta}{\sigma^2}}.$$

# Bibliography

- [Bau] F. Baudoin. Stochastic calculus lectures. <https://fabricebaudoin.wordpress.com/category/stochastic-calculus-lectures/>.
- [BM06] D. Brigo and F. Mercurio. Interest rate models - theory and practice. *Springer Finance*, Second Edition, 2006.
- [BR96] M. Baxter and A. Rennie. Financial calculus - an introduction to derivative pricing. *Cambridge University Press*, 1996.
- [BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.
- [CCM98] P. Carr, E. Chang, and D. Madan. The variance gamma process and option pricing. *European Finance Review*, 2:79–105, 1998.
- [CIR85] J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–408, Mar 1985.
- [CM99] P. Carr and D. Madan. Option valuation using the fast fourier transform. March 1999.
- [CT04] R. Cont and P. Tankov. Financial modelling with jump processes. *Chapman and Hall/CRC Financial Mathematics Series*, 2004.
- [Fel51] W. Feller. Two singular diffusion problems. *Annals of mathematics*, 54(1), July 1951.
- [FST10] M. Fujii, Y. Shimada, and A. Takahashi. A note on construction of multiple swap curves with and without collateral. *Financial Research and Training Center*, February 2010.
- [Hes93] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.
- [HP83] J.M. Harrison and R.S. Pliska. A stochastic calculus model of continuous trading: complete markets. *Stochastic processes and their applications*, 15:313–316, 1983.

- [HW90] J. Hull and A. White. Pricing interest-rate derivative securities. *The Review of Financial Studies*, 3(4):573–592, 1990.
- [Ito44] K. Ito. Stochastic integral. *Proc. Imp. Acad. Tokyo*, 20:514–524, 1944.
- [Ito51] K. Ito. On a formula concerning stochastic differentials. *Nagoya Math. Journal*, 3:55–65, 1951.
- [JS03] J. Jacod and A. Shiryaev. Limit theorems for stochastic processes. *Springer*, 2003.
- [Mik98] T. Mikosch. Elementary stochastic calculus. *World Scientific*, 6, 1998.
- [Pit10] V. Piterbarg. Funding beyond discounting: Collateral agreements and derivatives pricing. *ResearchGate*, January 2010.
- [Sat13] K. Sato. Lévy processes and infinitely divisible distributions. *Cambridge Studies in Advanced Mathematics*, 2013.
- [Sch03] W. Schoutens. Lévy processes in finance. *Wiley series in probability and statistics*, 2003.
- [Vas76] O. Vasicek. An equilibrium characterisation of the term structure. August 1976.
- [Veg74] F. De La Vega. On almost sure convergence of quadratic brownian variation. *The Annals of Probability*, 2(3):551–552, 1974.