Spread Options Pricing with Stochastic Correlation

by

Nikita Merkulov

Submitted to

Central European University

Department of Mathematics and Its Applications

In partial fulfillment of the requirements for the degree of Master of Science

Supervisor: Prof. László Márkus

Budapest, Hungary

2017

Acknowledgments

I would like to express my sincere gratitude to:

- my supervisor, Prof. László Márkus, for excellent lecture notes that helped me to get familiar with the basics of financial math, and for guidance during the work on the thesis;
- director of our MS program, Prof. Pál Hegedűs, for always being ready to give a good advice, for running the MS seminar, and, of course, for the great algebra lectures;
- my wife and my parents for support, invigoration and faith in me.

Abstract

We introduce an approximate formula for the price of a spread option, assuming that the two underlying assets are correlated via a Jacobi process. We also give an overview of the related theory: pure stochastic analysis, Black-Scholes-Merton ideas applied to the classical one-stock market and to the two-stocks market, and the extensions of the corresponding models.

Contents

Chapter 1

Introduction

The asset price modelling is important for those who wish to predict the change in market prices over time, advance our understanding of how market works, and, ultimately, to be able to make profit from the prices fluctuations. However, a model can only approximate the reality. The more advanced the model is, the higher the chances that it will reflect the real state of things well are.

A model is characterized by general expressions of how prices S evolve, e.g.

$$dS(t) = S(t)(\mu dt + \sigma dW(t))$$

(this is a *stochastic differential equation* discussed in the body of the thesis). Once these expressions are set - it is a challenging practical task to set them - one can take the prices known from real market history and *calibrate* the model, i.e. estimate the parameters the model depends on. The calibrated[1] model can be used to predict prices.

When we trust our model, our calibration methods and their predictions, the next logical step is to determine price of *derivatives*. Derivatives are financial instruments (simply, contracts) which price depend on other underlying assets (such as commodities, currencies, market indexes, stocks etc.)[2]. On the world market, thousands and millions of various derivatives - futures, swaps, options - are being traded every minute. Hence, millions of people are interested in derivatives pricing. Mathematics is here to suggest pricing that is reasonable, logical and fair[3].

The derivatives we focus on in the thesis are *spread options*. An option is a contract that gives its owner the right to buy or sell an asset for a specific price, at a specific date or within a specific period. In case of spread option, this asset is the difference (spread) between two other assets. Spread options are simple and widely used derivatives, yet there are still challenges to overcome regarding their pricing.

Of a particular interest when it comes to spread option pricing is dependance between the two assets. Let us illustrate it with an example of a *crack spread*[4], that is, a spread option written on the difference between crude and refined petroleum products - let it be crude oil and heating oil for definiteness. Prices of both depend on variety of hardly-predictable (stochastic) factors like laws, taxes, technologies. The two prices are obviously highly correlated - furthermore, the correlation coefficient itself can change over time. In

---

[1]Calibration is, of course, not a finished action but rather a process that needs to be performed when there is a new market data available.

[2]*Asset* is a financial term best defined by examples. Often, this word can be replaced with "something" (or "something that has a value") without losing much sense.

[3]"Fair" from the market perspective, human definitions of "fair" may differ.

[4]For more examples of spread options they use on real markets, see a wonderful review paper [5].

the thesis, we consider different approaches to correlation, starting from it being a constant and later developing a *stochastic correlation* model.

Modern financial theory would not be possible without mathematical tools that appeared in 1940-1960s: Itô integral [17] and formula [18], stochastic differential equations, Girsanov theorem [14]. All these go by name *stochastic analysis* (or stochastic calculus) and are considered in chapter 2. Chapter 3 is a brief review of the classical 1973's Black-Scholes-Merton approach to pricing [4], [30] and of its extensions. In chapter 4, we turn to the spread options. Again, we start from the 1978's Margrabe model [28] and then outline more recent ideas of [12]. Finally, in chapter 5 we (basing on methods presented in [26]) introduce a pricing formula for a spread option with stochastic correlation.

Chapter 2

Stochastic analysis

Price dynamic is usually modelled as a stochastic process. A convenient way to work with a stochastic process is to define it as a solution of a stochastic differential equation. To introduce stochastic differential equations, we first need a notion of stochastic integral, usually referred as Itô integral, for it was Kiyoshi Itô who pioneered the field in 1940-1950s.

Remark 2.0.1. Throughout the thesis we will assume, often without saying, that we are given a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. Recall that the filtration $\mathcal{F}_t$ is a monotonously increasing sequence of σ-algebras.

We will also use Φ to denote the cumulative distribution function of the standard normal distribution.

Consider a real-valued stochastic process $\{X_t\}_{t\in[0,T]}$ ($\forall t \ X_t : \Omega \to \mathbb{R}$ is a random variable). The index t will be interpreted as (continuous) time. Here, $T \leqslant \infty$, i.e. the time horizon may be either finite or infinite. Note that one can also think about the set $\{X_t\}$

as a function $X : [0, T] \to L^1(\Omega)$.[1] This is why the notations X_t and $X(t)$ are often used interchangeably in the context.

2.1 Itô integral

Let us outline how to construct the Itô integral. In this section, we follow the approach of [24], for Itô's own 1944's approach is somewhat overcomplicated due to the difficulties of the wartime.

2.1.1 Integral of simple processes

Definition 2.1.1. Let

$$X_t = \sum_{i=0}^{n-1} \alpha_i \mathbb{I}_{(t_i, t_{i+1}]}(t),$$

where $0 = t_0 < t_1 < \ldots < t_n = T$, $\forall i$ α_i is an F_{t_i}-measurable and square-integrable random variable, $\mathbb{I}_{(t_i, t_{i+1}]}$ is an indicator function. Then the process X is called a simple process.

In particular, a simple process is left-continuous and adapted to the filtration $\mathcal{F}_t$. Now we define Itô integral of a simple process with respect to a Wiener process in a very predictable and natural way.

Definition 2.1.2. Consider a simple process X and a Wiener process W. Then the Itô integral

$$I_T(X) = \int_0^T X_s dW_s := \sum_{i=0}^{n-1} \alpha_i (W_{t_{i+1}} - W_{t_i}).$$

---

[1] Here, $L^1(\Omega)$ is the space of all integrable real random variables. Of course, not all random variable are integrable, but they are usually assumed to be when it comes to financial applications. In fact, we will not consider anything more general than semimartingales (see below).

Note that $I_T(X)$ is a random variable. Also note that if we set $X_s^{(t)} := X_s \mathbb{I}_{[0,t]}(s)$ for some $t \in [0,T]$, then the process $X^{(t)}$ is again a simple one. Hence, we can define the integral with the upper limit different from T.

Definition 2.1.3. With the above notations,

$$I_t(X) = \int_0^t X_s dW_s := \int_0^T X_s^{(t)} dW_s = \int_0^T X_s \mathbb{I}_{[0,t]}(s).$$

One can see that the Itô integral is an integral of a stochastic process with respect to a stochastic process. It is also a stochastic process itself as a function of the upper limit.

Itô integral of a simple process (and a more general Itô integral considered below) has a few easy-to-prove properties: it is linear, additive, has continuous trajectories and is a martingale. Among the more important ones, there is an equality

$$E(\int_0^t X_s dW_s) = 0$$

and the so-called Itô isometry

$$E\left(\int_0^t X_s dW_s \int_0^t Y_s dW_s\right) = E\left(\int_0^t X_s Y_s ds\right).$$

2.1.2 Integral of square-integrable processes

Just as in pure functional analysis, every "good" process can be approximated by a sequence of simple processes.

Theorem 2.1.4. *Let Y be a measurable and $\mathcal{F}_t$-adapted process with*

$$E\left(\int_0^T (Y_s)^2 ds\right) < \infty.$$

Then there exists a sequence of simple processes $X^1, X^2, \ldots$ *satisfying*

$$E\left(\int_0^T (Y_s - X_s^n)^2 ds\right) \to 0,\ n \to \infty.$$

In other words, $X_s^n \to Y_s$ *in* $L^2([0,T] \times \Omega)$*.*

This theorem, together with Itô isometry, allows to define Itô integral of a square-integrable process.

Definition 2.1.5. With the notations of the theorem above,

$$\int_0^T Y_s dW_s := \lim_{n\to\infty} \int_0^T X_s^n dW_s.$$

For Itô integral of a square-integrable process, the above properties (linearity, continuity, Itô isometry etc.) hold.

2.1.3 Further extensions

In a similar way, if a process Z is square-integrable only with probability 1, then there is a sequence of square-integrable processes $Y^1, Y^2, \ldots$ such that

$$\int_0^T (Z_s - Y_s^n)^2 ds \to 0$$

in probability when $n \to \infty$. This fact allows to extend the notion of Itô integral to processes that are square-integrable with probability 1.

Without going into details too much, let us note that by this "limiting" approach an Itô integral of any semimartingale[2] can be defined. On the other hand, if we require the

[2] Semimartingale is a sum of a local martingale and an adapted process with locally bounded variation.

integral to continuously depend on the integrands, there is no way to extend the notion beyond the class of semimartingales. An important (for the Itô's formula we introduce below) fact about semimartingales is that they have a finite quadratic variation.

2.2 Itô's formulae and stochastic differential equations

2.2.1 Itô process

Definition 2.2.1. Let W be a Wiener process, $\mathcal{F}_t^W$ the filtration generated by W. Let a, b be $\mathcal{F}_t^W$-adapted processes satisfying

$$\int_0^T |a_s| ds < \infty, \int_0^T (b_s)^2 ds < \infty$$

with probability 1. Let X be a process satisfying

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t b_s dW_s.$$

Then X is called an Itô process.

The main reason to introduce the Itô process now is to make the reader familiar with the differential form of the definition of Itô process:

$$dX_t = a_t dt + b_t dW_t.$$

This is the shorthand usually used to write the Itô's formulae, stochastic differential equations, and, ultimately, pricing models.

Note that the shorthand also gives us a slightly tautological way to define the Itô integral (of a semimartingale Y) with respect to an Itô process X:

$$\int_0^T Y_s dX_s := \int_0^T Y_s a_s ds + \int_0^T Y_s b_s dW_s.$$

2.2.2 Itô's formulae

One of the most important differences between the real calculus and stochastic one is the change of variable formulae. They are again named after Kiyoshi Itô [18]. The simplest Itô formula is given as a theorem below.

Theorem 2.2.2. *Let f be a twice differentiable real function, W a Wiener process. Then*

$$df(W_t) = f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt.$$

A similar formula can be written for processes other than Wiener one.

Theorem 2.2.3. *Let f be a twice differentiable real function, X an integrable continuous martingale with quadratic variation $[X]_t < \infty$. Then*

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)d[X]_t.$$

The proofs of the theorems are based on the Tailor formula and the heuristic argument (which can in fact be proven rigorously) that $o((dX_t)^2) = o(dt)$ for processes with finite quadratic variation.

The latter Itô formula is frequently applied to Itô processes.

Theorem 2.2.4. *Let f be a twice differentiable real function, X an Itô process*

$$dX_t = a_t dt + b_t dW_t$$

(where a_t, b_t satisfy the conditions in the definition 2.2.1 of Itô process). Then

$$df(X_t) = (f'(X_t)a_t + \frac{1}{2}f''(X_t)b_t^2)dt + f'(X_t)b_t dW_t.$$

If the function is not time homogenous, another term appears.

Theorem 2.2.5. *Let $f = f(t, x)$ be a twice differentiable function, X an Itô process*

$$dX_t = a_t dt + b_t dW_t.$$

Then

$$df(t, X_t) = (f_t^{'}(t, X_t) + f_x^{'}(t, X_t)a_t + \frac{1}{2}f_{xx}^{''}(t, X_t)b_t^2)dt + f_x^{'}(t, X_t)b_t dW_t.$$

More generally:

Theorem 2.2.6. *Let $f = f(t, x^1, \ldots, x^d)$ be a twice differentiable function, $X = (X^1, \ldots, X^d)$ an vector-valued integrable continuous martingale with quadratic variations $[X^i, X^j]_t < \infty \quad \forall i, j \in \{1, \ldots, d\}$. Then*

$$df(t, X_t) = f_t^{'}(t, X_t)dt + \sum_{i=1}^{d} f_{x^i}^{'}(t, X_t)dX_t^i + \frac{1}{2}\sum_{i,j=1}^{d} f_{x^i,x^j}^{''}(t, X_t)d[X^i, X^j]_t.$$

2.2.3 Stochastic differential equations

The theory of stochastic differential equations is in fact independent of financial mathematics. An interested reader is referred to, for example, [31].

Let W be a Wiener process, $\mathcal{F}_t^W$ the filtration generated by W. The equation

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \tag{2.1}$$

where $a(t, x)$, $b(t, x)$ are Borel-measurable real functions, is called a stochastic differential equation - SDE (on a process X). An SDE usually comes with an initial condition

$$X_0 = \xi, \tag{2.2}$$

where ξ is $\mathcal{F}^W_\infty$-independent square-integrable random variable.

Strong solution

Definition 2.2.7. The strong solution of the equation (2.1) given this initial condition (2.2) is an $\mathcal{F}^W_t \vee \sigma(\xi)$-adapted continuous stochastic process X satisfying

$$X_t = \xi + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s$$

and

$$\int_0^t (|a(s, X_s)| + b^2(s, X_s))ds < \infty$$

for every t with probability 1.

As one would expect, there is a existence and uniqueness theorem for the strong solution of an SDE.

Theorem 2.2.8. *Suppose the coefficients in (2.1) are Lipshitz-continuous linearly growing functions:*

$$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| < C|x - y|,$$

$$|a(t, x)| + |b(t, x)| < K(1 + x^2)$$

for some positive constants C and K and $\forall t$. Then there exists X - a strong solution of (2.1) given (2.2). This solution is unique, i.e. $P\{X_t = \hat{X}_t \ \forall t\} = 1$ for a strong solution $\hat{X}$. Furthermore,

$$E \int_0^t (X_s)^2 ds < \infty \ \forall t < \infty.$$

This theorem straightforwardly extends to multidimensional SDEs with vector-valued solution process X. The Lipshitz condition can in fact be relaxed - see, for example, [40].

Weak solution

One can think about a strong solution as follows: we are given the Wiener process W, and then we construct a process X that is $\mathcal{F}_t^W$-adapted (and satisfies the SDE). On the contrary, if we only have the coefficients a, b of SDE, we construct a "proper" Wiener process W along with X in order to get a so-called weak solution.

Definition 2.2.9. The weak solution of the SDE (2.1) under the initial condition

$$X_0 \sim Q_\xi$$

is a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ and a pair of processes (X, W) such that

- X is a $\mathcal{F}_t$-adapted process with continuous trajectories;
- W is a $\mathcal{F}_t$-adapted Wiener process;
- X satisfies (2.1) in a usual sense with the particular W;
- $X_0 \sim Q_\xi$.

As the names suggest, a strong solution is a weak solution, but the converse is not always true. A famous counterexample is known as Tanaka equation

$$dX_t = sign(X_t)dW_t,\ X_0 = 0.$$

It has no strong solution, although a weak solution is easy to construct.

Theorem 2.2.10 (Skorokhod, Stroock, Varadhan)**.** *Suppose the coefficients of the stochastic differential equation (2.1) are bounded and continuous functions, and the* $2+\varepsilon$ *moment* *(for some $\varepsilon > 0$) of the distribution Q_ξ is finite. Then the SDE has a weak solution under the initial condition*

$$X_0 \sim Q_\xi.$$

When one transforms an SDE looking for a weak solution, she needs to follow what is a Wiener process and what is not. Hence the interest in the relations between different Wiener processes. A powerful mathematical tool to treat these issues is considered briefly in the next section.

2.3 Measure change theory

Definition 2.3.1. Let X be an Itô process defined in 2.2.1, $\mathcal{F}_t^X$ the filtration generated by X. X is called a diffusion process, if the coefficients a, b are $\mathcal{F}_t^X$-adapted.

Trajectories of a diffusion process are continuous functions, so it makes sense to speak about the distribution of a diffusion process, that is, a corresponding measure on the space $C[0,T]$ with the Borel σ-algebra. Let μ_X denote the distribution of a diffusion process X.

In the next few statements, X is a diffusion process with diffusion coefficient 1:

$$dX_t = a_t dt + dW_t. \tag{2.3}$$

Lemma 2.3.2. *Let the process a in (2.3) satisfy*

$$P\left(\int_0^T a_t^2 dt < \infty\right) = 1,$$

$$E\exp\left(-\int_0^T a_t dW_t - \frac{1}{2}\int_0^T a_t^2 dt\right) = 1.$$

Then $\mu_X \ll \mu_W$ and $\mu_W \ll \mu_X$*, with the Radon-Nikodim derivative given by*

$$\frac{d\mu_W}{d\mu_X} = E\left(\exp\left(-\int_0^T a_t dX_t + \frac{1}{2}\int_0^T a_t^2 dt\right) |\mathcal{F}_T^X\right).$$

In an important special case, the formula simplifies.

Lemma 2.3.3. *Let* a_t *in (2.3) be* $\mathcal{F}_T^X$*-measurable* $\forall t$*. Then* $\mu_X \ll \mu_W$ and $\mu_W \ll \mu_X$*, with the Radon-Nikodim derivative given by*

$$\frac{d\mu_W}{d\mu_X} = \exp\left(-\int_0^T a_t dX_t + \frac{1}{2}\int_0^T a_t^2 dt\right).$$

Remark 2.3.4. $\mathcal{F}_T^X$-measurability of a_t means that a depends on randomness only through X, i.e. $a_t(\omega) = a_t(X_t(\omega))$.

Remark 2.3.5. X_t is not an $\mathcal{F}_t$-martingale unless $a_t \equiv 0$. However, if we define

$$M_t = \exp\left(-\int_0^t a_s dW_s - \frac{1}{2}\int_0^t a_s^2 ds\right)$$

and

$$Y_t = M_t X_t,$$

then Y_t is an $\mathcal{F}_t$-martingale.

The latter statement is a special case of the Girsanov Theorem [14].

Theorem 2.3.6 (Girsanov)**.** *Let* W *be a Wiener process,* a *an* $\mathcal{F}_t^W$*-adapted process with* $\int_0^T a_t^2 dt < \infty$ *with probability 1. Suppose that*

$$Z_t := \exp\left(\int_0^t a_s dW_s - \frac{1}{2}\int_0^t a_s^2 ds\right)$$

is an $\mathcal{F}_t^W$-martingale under the P measure. Define a new measure Q_a as

$$dQ_a(\omega) = Z_T(\omega)dP.$$

Then the process $\tilde{W}_t$ defined as $\tilde{W}_t = W_t + \int_0^t a_s ds$ under the Q_a measure.

Remark 2.3.7. Under the P measure, the process $\tilde{W}$ is nothing else but the same diffusion process with diffusion coefficient 1, also known as a drifting Wiener process.

The martingale property of Z is not always easy to check, but with some restrictions on a it holds. These restrictions are known as Novikov [25] and Kazamaki [25] conditions. They are defined for any X local martingale on $[0, T]$ ($T \leqslant \infty$) with continuous trajectories.

Definition 2.3.8. We say that X satisfies the Novikov condition, if

$$E \exp\left(\frac{1}{2}[X]_t\right) < \infty \ \forall t > 0.$$

Definition 2.3.9. We say that X satisfies the Kazamaki condition, if

$$\sup_t E \exp\left(\frac{1}{2}X_t\right) < \infty.$$

Remark 2.3.10.

- If $\exp\left(\frac{1}{2}X_t\right)$ is a uniformly integrable submartingale, the Kazamaki condition holds.
- Novikov condition follows from Kazamaki condition.
- If X is uniformly integrable (sub)martingale, the Kazamaki condition simplifies to $E \exp\left(\frac{1}{2}X_T\right) < \infty$.

There are modifications of the Girsanov theorem that allow to get the Wiener process from diffusion processes with an arbitrary diffusion coefficient. Furthermore, a Wiener process is, of course, not the only way to model the randomness. We will see that some pricing models put other processes under the differential in the SDE. So there are generalisations of the Girsanov theorem that describe the measure and process change needed to get a new-measure-local martingale from an old-measure-local martingale. We will, however, not cover them here for it would require a set of new definitions.

Chapter 3

Pricing models with one stock

In this chapter, we show how mathematical concepts considered before are applied to finance. The whole story began in 1973, after works of Black and Scholes [4] and Merton [30]. We will cover the most fundamental ideas, as well as some ways to generalize and improve them. Many explanations in the chapter are quoted by [29].

The basis of modern economic life is the company owned by its shareholders; the shares provide partial ownership of the company, pro rata with investment. Shares are issued by companies to raise funds. They have value, reflecting both the value of the company's real assets and the earning power of the company's dividends. *Stock* is the generic term for assets held in the form of shares. With publicly quoted companies, shares are quoted and traded on a stock exchange.

3.1 Classical Black-Scholes approach

3.1.1 The model; derivative securities

Let us start with a market where only two assets are traded: a risk free bank account, which is usually called a bond, and a stock ("defined" above). The market is described by the bond and the stock prices. We will model the price processes of both bond and stock by continuous time processes. In practice, it is impossible to observe stock prices varying continuously in time, and the prices themselves are discrete. Nevertheless, the continuous-variable, continuous-time process is proved to be a useful model for many purposes. Throughout, we denote the bond price process by B and the stock price process by S.

Let B satisfy an ordinary differential equation

$$dB(t) = rB(t)dt$$

with a constant r. Its solution is

$$B(t) = B(0)e^{rt}.$$

Naturally, the bank account value simply grows exponentially over time.

To model the stock price, we will employ stochastic differential equations introduced in the previous chapter:

$$dS(t) = S(t)(\mu dt + \sigma dW(t)).$$

For now, let μ and σ be constants. In this model, as well as in more complicated non-constant ones, μ is called *drift* and σ is called *volatility*.

Shorty after applying the Itô's formula, one can get the (strong) solution of this SDE as

$$S(t) = S(0) \exp\left(\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W(t)\right).$$

Such process S is called Geometric Brownian Motion[1]. This is a process with continuous trajectories, it has lognormal marginal distribution. Note that the growth *proportions* of the process have independent increments.

Remark 3.1.1.

$$\ln(S(t)) = \ln(S(0)) + \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W(t),$$

i.e. the process $\ln(S)$ is just a drifting Wiener process.

Once we know the processes S and B, we can derive the price[2] of *derivative securities*. A derivative security (or derivative for short, or contingent claim) is a security whose value depends on the value(s) of other more basic underlying securities. To be more precise, it is a financial contract whose value at expiration date T is determined exactly by the price process of the underlying financial assets (or instruments, typically the stock price) up to time T.

A popular derivative to work with is a *European (vanilla) option*. European options are contracts that give the owner the right (not the obligation), to buy (*call* option) or sell (*put* option) the underlying asset at a prespecified price K (*strike price*), on the option's expiration date T (*maturity*). Thus, the price of a European option is determined by the price process of the underlying assets *at* the time T only. More general derivatives with

---

[1] We do not use the term "Brownian Motion" when we speak about a Wiener process in order to avoid confusion with the physical meaning. Similarly, mathematicians could name S a "Geometric Wiener process". However, the latter term is almost non-existent in the literature.

[2] More precisely, the fair price, not formally defined here. Intuitively, a fair price is the only reasonable price within the framework of a given model. For a proper definition, including the concept of completely admissible self-financial strategy, see e.g. [20].

this property are called European-type derivatives. Their payoff can be written as $f(S(T))$. For example, $f(x) = (x-K)_+$[3] for a European call option and $f(x) = (K-x)_+$ for a European put option.

3.1.2 Equivalent martingale measure and the price

We now need an equivalent martingale measure (or for several) - that is, a measure Q_μ such that:

- $Q_\mu \sim P$;
- the discounted stock price $\tilde{S} := e^{-rt}S(t)$ is a local martingale with respect to Q_μ.

An equivalent martingale measure is also known as *risk-neutral* measure.

An integral with respect to the Wiener process is a local martingale, so it is enough to give a representation of the discounted stock price as a Wiener-integral. To achieve this, we change the probability measure by Girsanov's theorem 2.3.6. Let $a(t) \equiv \frac{\mu-r}{\sigma}$, then define

$$dQ_\mu = \exp\left(\int_0^T a_s dW_s - \frac{1}{2}\int_0^T a_s^2 ds\right) dP = \exp\left(\frac{\mu-r}{\sigma}W(T) - \frac{1}{2}(\frac{\mu-r}{\sigma})^2 T\right) dP.$$

With respect to Q_μ, the process

$$W_\mu := \frac{\mu-r}{\sigma}t + W(t)$$

is a standard Wiener process. On the other hand,

$$d\tilde{S}(t) = \tilde{S}(t)\sigma dW_\mu, \tag{3.1}$$

[3] y_+ is the notation for $\max(y, 0)$.

i.e. $\tilde{S}$ is a local martingale with respect to Q_μ.

Remark 3.1.2. In general, there are conditions to be imposed on the model to ensure that a risk-neutral measure exists. The strongest results in this direction is known as "no free lunch with vanishing risk" (NFLVR). The reader is referred to [23], [9] for the details of this theory.

The existence of a risk-neutral measure allows one to compute the prices of derivatives.

Theorem 3.1.3. *Within the framework of the Black-Scholes model, the fair price of the European-type derivative with the payoff $f(S(T))$ is the expectation $E_\mu(e^{-rT}f(S(T)))$, taken under the measure Q_μ.*

The expectation with respect to risk-neutral measure is called the risk-neutral expectation for short; similarly, one can talk about risk-neutral pricing. There are, of course, similar theorems for non-European derivatives, but they all operate with this fundamental concept of risk-neutral.

3.1.3 Black-Scholes formula

The assumption that both drift and volatility are constants is a huge oversimplification, and, as we will see, the Black-Scholes model can be improved in numerous ways. Nevertheless, there are reasons to like it as it is. One of the most important ones is the closed formula for the option price.

Theorem 3.1.4 (Black-Scholes formula)**.** *Within the framework of the Black-Scholes model, the fair price of a European call option with the strike K and maturity T at the time 0 is*

$$C(S,0) = S(0)\Phi(d_2) - Ke^{-rT}\Phi(d_1),$$

where[4]

$$d_1 = \frac{\ln(S(0)/K) + Tr}{\sigma\sqrt{T}} - \frac{1}{2}\sigma\sqrt{T},\ d_2 = \frac{\ln(S(0)/K) + Tr}{\sigma\sqrt{T}} + \frac{1}{2}\sigma\sqrt{T}.$$

The fair price of a European put option with the same parameters is

$$P(S, 0) = Ke^{-rT}\Phi(-d_1) - S(0)\Phi(-d_2).$$

The proof is based simply on a substitution of the payoff $f(x) = (x - K)_+$ or $f(x) = (K - x)_+$ to the expectation in the theorem 3.1.3.

Note that μ is not presented in the formulae. Roughly speaking, we eliminate the drift when switching to the risk-neutral measure (see (3.1). Therefore, in the next section we mostly play with the volatility but not with the drift.

3.2 Extending the model

The issue with the Black-Scholes model is a considerable gap between the prices it suggests and the real market prices. There is an immense number of somewhat better continuous-time trading models, and there are hundreds of interesting and applicable results. This is why this section mostly consists of references.

Before we introduce some real model extensions, let us mention one that came up in the same year with the original model. If the interest rate and the volatility are not constant but depend on time ($r = r(t)$, $\sigma = \sigma(t)$), almost nothing changes, and an expression for the price similar to classical Black-Scholes formula can still be derived. One should only put the averaged parameter instead of the constant: $\frac{1}{T}\int_0^T r(t)dt$ instead of r and $\frac{1}{T}\int_0^T \sigma(t)dt$ instead of σ. See [30].

---

[4] Throughout the thesis, Φ is the cumulative distribution function of the standard normal distribution.

3.2.1 Volatility models

If we try to estimate the constant σ in the Black-Scholes formula from the real market data, we get a so-called implied volatility - a function of strike K and maturity T. Its empirical properties (e.g. a famous "smile" - U-shape dependance on the strike price) are desired properties of volatility in non-constant volatility models.

In *local volatility* models, we are given a deterministic function $\sigma(t, s)$ and the stock price as

$$dS(t) = S(t)(\mu dt + \sigma(t, S(t))dW(t)).$$

The question is the choice of $\sigma(t, s)$. In 1975, Cox and Ross [7] proposed the model with Constant Elasticity of Variance: $\sigma(t, S(t)) = \sigma S^{\beta}(t)$, where σ on the right-hand side and β are constants. However, the more constants we keep in the model for simplicity, the further it is from the reality. There is a way to get the function $\sigma(t, s)$ directly from the observed option prices [10]. On the other hand, if one has a lot of data for options with different strikes and maturities, one can get the local volatility function for a given strike and maturity as a solution of a partial differential equation [11].

Local volatility models do not produce the proper dynamics of the implied volatility surface, i.e. the parameters of equations calibrated at different times may well be entirely different. There are other properties that conflict with empirical evidence, such as perfect correlation with the stock price and lack of volatility clustering. To overcome these drawbacks, one can use a *stochastic volatility* model instead. The local volatility is, of course, already stochastic, but only through the stock price. In stochastic volatility models the volatility is driven by its own SDE.

A stochastic volatility model can usually be written as

$$\begin{cases} dS(t) = S(t)(\mu dt + \sigma(t)dW(t)) \\ \sigma(t) = f(Y(t)) \\ dY(t) = \lambda(\eta - Y(t))dt + g(t, Y(t))d\widehat{W}(t)) \end{cases},$$

where η, λ are constants and f, g are deterministic functions. η is the mean of $Y(t)$, and λ is a rate of mean reversion - indeed, a well-known empirical fact is that the volatility is mean-reverting.

Remark 3.2.1. There is a not-yet-mentioned parameter in this system. $\widehat{W}$ is, of course, a Wiener process, but it is not necessarily independent from W. In general, they are correlated with the correlation coefficient $\rho \in [-1, 1]$. It means that $\widehat{W}$ can be written as

$$\widehat{W}(t) = \rho W(t) + \sqrt{1 - \rho^2}\widetilde{W}(t),$$

with a further $\widetilde{W}$ process independent of W.

For now, let ρ be constant. In fact, in the next few models $\rho = 0$.

In Hull-White model [16], Y is a Geometric Brownian Motion ($\eta = 0$, $\lambda = -1$, $g(t, y) = Cy$), meaning it is rather simple to calibrate, but lacks the mean-reversion property. Scott [35] employs the (mean-reverting) Gaussian Ornstein-Uhlenbeck process ($\eta, \lambda \neq 0$, $g(t, y) \equiv \beta$). However, it can take negative values, so a transformation $f(y) = e^y$ is needed for the volatility to be positive, which affects the tractability of the model.

The shortcomings of these models made the Cox-Ingelsoll-Ross (CIR) [8] process a popular choice for Y. CIR process is both mean-reverting and non-negative. In both Ball-Roma [1] and Heston [15] models Y is a CIR process. The significant difference between them is that Heston takes the dependance between price and volatility into account: $\rho \neq 0$.

It allows[5] the model to replicate the phenomenon known as *leverage effect*: the negative shocks in the asset price are often followed by positive shocks in volatility. Therefore, often in the applications $\rho < 0$, while there is, of course, no mathematical reasoning for this limitation.

3.2.2 Random jumps models[6]

So far, the stock prices were changing continuously, driven by a Wiener process (under the exponent). However, prices on the real markets can change rapidly (jump) when some new information reaches the market. The mathematical framework covering these situations is set by Lévy processes.

Definition 3.2.2. A stochastic process $X(t)$ on $[0, T]$ is called a Lévy process, if it

- has independent increments, that is,

$$X(t_0), X(t_1) - X(t_0), \ldots, X(t_n) - X(t_{n-1})$$

  are independent for any $t_0 < t_1 < \cdots < t_n \in [0, T]$;

- has stationary increments, that is,

$$X(t + h) - X(t) \sim X(s + h) - X(s) \ \forall t, s, t + h, s + h \in [0, T];$$

[5] On the other hand, as mentioned in [1], this complication makes the model less tractable. We are not trying to prove that some models are better than the others, but rather give the reader an idea about their strengths and weaknesses.

[6] We recommend a good practical overview of this issue [38], or more detailed papers [29], [33].

- is continuous in probability, that is,

$$\forall t \in [0,T], \epsilon > 0 \lim_{h \to 0} P(|X(t+h) - X(t)| \geq \epsilon) = 0;$$

- there exists a modification with càdlàg (everywhere right-continuous, and with left limits everywhere) trajectories.

To help the reader understand what Lévy processes are, we give an informal version of the important theorem. For details on the statement and proof see e.g. [32].

Theorem 3.2.3 (Lévy-Itô decomposition). *Let X be a "good-enough" Lévy process with values in $\mathbb{R}$. Then there exists a constants γ, σ and a standard Wiener process W such that $\forall t$*

$$X(t) = \gamma t + \sigma W(t) + Y(t),$$

where Y is a pure jump process.

In other words, a Lévy process is a would-be diffusion process that also make random jumps at random times. A degenerate examples of Lévy processes that only make zero jumps are diffusion processes and Wiener process.

Definition 3.2.4. A Lévy process is called a subordinator, if its trajectories are non-decreasing with probability 1.

Subordinators enable us to make "snapshots" of another Lévy process - in particular, a Wiener process - at random times.

Definition 3.2.5. Let S be a subordinator, W a Wiener process independent of S, μ, σ constants. Then the process

$$X(t) = \mu S(t) + \sigma W(S(t))$$

is called a subordinated Brownian Motion.

The subordinated Brownian Motion is a Lévy process with jumps. In this model we look at a drifting Brownian Motion in a new stochastic time scale. This time scale can be regarded as "business time", that is, the sequence of random times of information "arrivals". Arriving information impulses the traders to enter transactions and instantaneously change the price of an asset.

This subordinating technique allows us to "get" useful processes "from" a Wiener process. The most famous processes obtained this way go under the names of variance gamma (VG) process [27], normal inverse Gaussian (NIG) process [2], or, more generally, Carr-Geman-Madan-Yor (CGMY) process [6]. After taking an exponent for non-negativity, these processes can be used to model randomness in asset prices. *Exponential Lévy* is the general prefix for these models.

Apart of the two mentioned parameters that enable to improve a one-stock pricing model - volatility and the stochastic process driving the price - there is also the interest rate r. Although the interest rate models and their effect on the pricing model are not covered in the thesis, they provide an interesting separate line of research. For more information, see e.g. [22] and references therein.

Chapter 4

Spread option pricing

The theory and models discussed in the previous chapter allow to price derivatives which depend on one stock. However, when a derivative price depend on two or more underlying prices, the situation may complicate. *Spread options* are among the derivatives that still remain without absolutely efficient and reliable pricing methodologies. Some explanations in the chapter are quoted by [5].

A spread option is an option written on the difference of two underlying assets, whose values at time t we denote by $S_1(t)$ and $S_2(t)$. We consider only spread options of the European type for which the owner has the right to be paid, at the maturity date T, the difference $S_2(T) - S_1(T)$, known as the *spread*. To exercise the option (i.e. to realize the right), the buyer must pay at maturity a prespecified strike price K. In other words, the payoff $f(S_1(T), S_2(T))$ of a spread option at maturity T is $(S_2(T) - S_1(T) - K)_+$.

We focus our attention on the particular case of a spread option with strike $K = 0$, that are also called *exchange options*. Besides the fact that the case $K = 0$, as we will see, can lead to a solution in closed form, it has also a practical appeal to the market participants. Indeed, it can be viewed as an option to exchange one product for another at no cost. In order to illustrate this fact, let us suppose that we want to buy one of two stocks S_1 or S_2, that we are indifferent to which one we own, but that at the end of a six-month period T we would, naturally, like to be in the position of someone who owns the better performing stock of the two. Obviously, we cannot tell which one will perform better over the next six months. We finally decided that S_1 is a better bet, but we are not quite sure of our pick, so we buy an option to exchange S_1 for S_2 in case S_2 outperforms S_1 after six months. Indeed, if the second stock ends up being the more valuable, i.e., if $S_2(T) > S_1(T)$, then the payoff $S_2(T) - S_1(T)$ of the option will exactly compensate us for our wrong choice. Short of the premium (i.e., the price we have to pay to own the option), our investment is the better of the two stocks over the next six months.

4.1 Classical Margrabe approach

Consider the market where a bond (see section 3.1.1)

$$B(t) = B(0)e^{rt}$$

and two other assets are traded. The two assets are assumed to be interdependent - say, gas and oil. More precisely, let the two prices be driven by stochastic differential equations

$$\begin{cases} dS_1(t) = S_1(t)(\mu_1 dt + \sigma_1 dW_1(t)) \\ dS_2(t) = S_2(t)(\mu_2 dt + \sigma_2 dW_2(t)) \end{cases}, \tag{4.1}$$

where

$$dW_1(t)dW_2(t) = \rho dt,$$

i.e. the Wiener processes W_1 and W_2 are correlated with the correlation coefficient ρ (we encountered correlated Wiener processes before - see remark 3.2.1). For now, let r,

$\mu_1 = \mu_2$[1], σ_1, σ_2 and ρ be constants[2]. This is known as Margrabe [28] model[3]. Its advantage is a closed formula for the price of a spread option.

Theorem 4.1.1 (Margrabe formula)**.** *Within the framework of the Margrabe model, the fair price of a European spread option with the strike* $K = 0$ *and maturity* T *at the time* 0 *is*

$$C(S_1, S_2, 0) = S_2(0)\Phi(d_2) - S_1(0)\Phi(d_1),$$

where

$$d_1 = \frac{\ln(S_2(0)/S_1(0))}{\sigma_0\sqrt{T}} - \frac{1}{2}\sigma_0\sqrt{T},\ d_2 = \frac{\ln(S_2(0)/S_1(0))}{\sigma_0\sqrt{T}} + \frac{1}{2}\sigma_0\sqrt{T}$$

and

$$\sigma_0 = \sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}.$$

The proof is based on the fact that, again, there exists a risk-neutral measure Q_{μ_1,μ_2} (see section 3.1.2), and the price of the spread option is given by the risk-neutral expectation $E_{\mu_1,\mu_2}(e^{-rT}(S_2(T) - S_1(T))_+)$. After this, the straightforward computation of the measure and the expectation gives the formula.

---

[1]The model straightforwardly extends to the case when drifts are not equal, but we will keep this assumption for simplicity.

[2]One can easily see the similarity with the Black-Scholes approach. This is why in the chapter 5 the price derived by Margrabe will eventually be called the "Black-Scholes" price.

[3]A fair name for the model would be Fisher-Margrabe model, for Fisher [13] reached similar conclusions at the same time. In fact, Fisher's results were published *in the same issue* of Journal of Finance with Margrabe's results. We would like to thank the author of [34] for putting this information into his paper.

4.2 Extending the model

Margrabe formula was introduced in 1978. Thanks to the mathematicians' efforts over almost four decades, two-stocks pricing models got all the improvements that we mentioned in section 3.2. In fact, both stochastic volatility and exponential Lévy model extensions are provided by the general theory introduced in [12]. We will briefly overview the results here, using also [3].

Consider a two-dimensional semimartingale $H = (H_1, H_2)$ with the canonical representation (see II.2.34 in [19]) of the form

$$H(t) = \int_0^t b(s)ds + \int_0^t c^{\frac{1}{2}}(s)dW(s) + \int_0^t \int_{\mathbb{R}^2} h(z)(\mu - \nu)(ds, dz) + \int_0^t \int_{\mathbb{R}^2} (z - h(z))\mu(ds, dz),$$

where the "drift" $b(t) = b(\omega, t) \in \mathbb{R}^2$, the "continuous volatility" $c(t) = c(\omega, t) \in \mathbb{R}^{2\times2}$, and μ, ν are measures associated with the process H - respectively, the jump-measure (a compound Poisson process here) and its compensator (see [33]). In a sense, this is a generalisation of Lévy-Itô decomposition (theorem 3.2.3).

Remark 4.2.1. Here, $W = (W_1, W_2)$ is a standard two-dimensional Wiener process. In the model (4.1), it could be possible if ρ is a constant or a deterministic function. However, if ρ is a stochastic process itself, W is not a two-dimensional Wiener process but rather just a vector of two Wiener processes. See chapter 5 for things that work in this case.

Let the price process $S = (S_1, S_2)$ be given by the stochastic exponential of H (see [36]):

$$S(t) = S(0)(\mathcal{E}(H))(t).$$

It turns out that if the price is defined this way, stochastic volatility and exponential Lévy models come up as a special case.

Up to several assumptions not mentioned here, the price of a spread option is given by a *one-dimensional* expectation with respect to so-called dual measure.

Theorem 4.2.2. *Within the framework of the above model, the fair price of a European spread option with the strike $K = 0$ and maturity T at the time 0 is*

$$C(S_1, S_2, 0) = S_1(0) E_{\widetilde{P}} \left(\exp\left(\int_0^T (r(s) - a(s)) ds \right) \left(1 - \frac{S_2(T)}{S_1(T)} \right)_+ \right),$$

where $r = r(t)$ is the interest rate[4]*, and both the function $a = a(t)$ and the dual measure $\widetilde{P}$ are defined in [3].*

We will not go into more details; this section is intended only to give the reader an idea about how to extend the model beyond classical approach. Instead, we turn out attention to the empirically justified case of stochastically correlated price-driving processes.

[4] As always, we assume the bond being traded along with the stocks.

Chapter 5

Extending the model - stochastic correlation

As before, consider two prices driven by stochastic differential equations

$$\begin{cases} dS_1(t) = S_1(t)(\mu_1 dt + \sigma_1 dW_1(t)) \\ dS_2(t) = S_2(t)(\mu_2 dt + \sigma_2 dW_2(t)) \end{cases}, \tag{5.1}$$

where

$$dW_1(t)dW_2(t) = \rho dt,$$

i.e. the Wiener processes W_1 and W_2 are correlated with the correlation coefficient ρ.

We have seen before that drifts and volatilities are not necessarily constant. In the previous chapter, the correlation coefficient could as well be time- and/or price-dependant. However, just as with volatility, empirical evidence suggests that on the real-world markets the correlation coefficient changes stochastically over time. As already hinted in remark 4.2.1, this case requires new methods.

The correlation should never be greater than 1 or smaller than -1. This is why the process that usually models the stochastic correlation is a Jacobi process.

Definition 5.0.1. The process ρ that satisfies

$$d\rho(t) = (\bar{\rho} - \beta\rho(t))dt + \sigma\sqrt{(h - \rho(t))(\rho(t) - f)}dW$$

(where W is, as always, a Wiener process) is called a Jacobi process.

Remark 5.0.2. For a Jacobi process ρ, $f \leqslant \rho(t) \leqslant h$ $\forall t$ almost surely, if the following holds for the *constants* $\bar{\rho}$, h, f and σ:

$$f < \bar{\rho} < h; \bar{\rho} - \beta f > \frac{\sigma^2(h-f)}{2}; \beta h - \bar{\rho} > \frac{\sigma^2(h-f)}{2}. \tag{5.2}$$

Since we are going to use the Jacobi process to model the correlation coefficient, we take f and h such that

$$-1 \leqslant f \leqslant h \leqslant 1. \tag{5.3}$$

Let us note that this is not the only possible approach. For example, the authors of [39] additionally employ Ornstein-Uhlenbeck process to model the correlation. However, the computations with the Ornstein-Uhlenbeck process in this role are not absolutely rigorous, for this process can leave the $[-1, 1]$ interval with positive probability. Therefore, we focus on the Jacobi process.

5.1 Black-Scholes-type pricing with stochastic correlation

[26] is devoted to pricing of quanto-options with stochastic correlation. Their price depend on two (stochastically correlated) asset prices, and this is the similarity between them and

the spread options we consider in the thesis. Despite the fact that the payoff functions are, of course, different, the methods presented in [26] can be applied to spread option pricing. Whenever the reader feels stuck in this chapter, she is encouraged to check [26] and the references therein for details.

Let ρ in our model for S_1 and S_2 (5.1) be a Jacobi process. We could additionally assume that there is a non-zero correlation between the price processes and the correlation process itself:

$$\begin{cases} dW_1(t)dW(t) = \rho_1 dt \\ dW_2(t)dW(t) = \rho_2 dt \end{cases},$$

However, the author of [26] empirically argues that it is safe to simplify the problem and take $\rho_1 = \rho_2 = 0$. We follow this approach and only consider W independent from W_1 and W_2.

We begin our analysis with only the correlation being stochastic but the drifts and volatilities being constant.

Let $C = C(S_1, S_2, t)$ be the price of the spread option. Then we can apply Itô's formula to get

$$dC = \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma_1^2 S_1^2 \frac{\partial^2 C}{\partial S_1^2} + \rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 C}{\partial S_1 \partial S_2} + \frac{1}{2}\sigma^2(h-\rho)(\rho-f)\frac{\partial^2 C}{\partial \rho^2}\right)dt +$$

$$\frac{\partial C}{\partial S_1}dS_1 + \frac{\partial C}{\partial S_2}dS_2 + \frac{\partial C}{\partial \rho}d\rho$$

(the arguments are omitted here and after). We now skip few technical steps one can find in [26] and simply state that the *risk-neutral* price in our case is given by the equation

$$\frac{\partial C}{\partial t} + (r_1 - r_2)S_1 \frac{\partial C}{\partial S_1} + (r_1 - \rho\sigma_1\sigma_2)S_1 \frac{\partial C}{\partial S_1} + (\bar{\rho} - \beta\rho)\frac{\partial C}{\partial \rho} +$$

$$\frac{1}{2}\sigma_1^2 S_1^2 \frac{\partial^2 C}{\partial S_1^2} + \frac{1}{2}\sigma_2^2 S_2^2 \frac{\partial^2 C}{\partial S_2^2} + \rho\sigma_1\sigma_2 S_1 S_2 \frac{\partial^2 C}{\partial S_1 \partial S_2} +$$

$$\frac{1}{2}\sigma^2 (h-\rho)(\rho - f)\frac{\partial^2 C}{\partial \rho^2} - r_1 C = 0,$$

where r_1, r_2 are interest rates.

The option payoff is $C(S_1, S_2, T) = (S_1(T) - S_2(T) - K)_+$. One can solve the above equation numerically, using this payoff as a terminal condition and a Monte-Carlo method to sample ρ. On the other hand, there is a highly non-trivial way to get a closed form approximation for the price C.

5.2 Closed form approximation for the price

As before, the reader can consult [26] for the computations that allow to get the results of subsections 5.2.1 and 5.2.2.

5.2.1 Probability kernel for the Jacobi process

In our notations, the Markov generator of the Jacobi process ρ is

$$H = \frac{\sigma}{2}(h-\rho)(\rho-\beta)\frac{d^2}{d\rho^2} + (\bar{\rho} - \beta\rho)\frac{d}{d\rho}.$$

Using the technique of [37], one can solve the eigenproblem

$$H\psi_n(\rho) = \lambda_n \psi_n(\rho).$$

The solutions are

$$\lambda_n = -\frac{\sigma}{2} n (n + \frac{2\beta}{\sigma^2} - 1)$$

and

$$\psi_n(\rho') = \left(\frac{(2n + \frac{2\beta}{\sigma^2} - 1)\Gamma(n + \frac{2\beta}{\sigma^2} - 1)n!}{2^{\frac{2\beta}{\sigma^2}-1}\Gamma(n + \frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)})\Gamma(n + \frac{2\bar{\rho} - 2\beta f}{\sigma^2(h-f)})} \right)^{\frac{1}{2}} P_n^{(\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}-1, \frac{2\bar{\rho}-2\beta f}{\sigma^2(h-f)}-1)}(\rho'),$$

where P_n are Jacobi polynomials given by

$$P_n^{(\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}-1, \frac{2\bar{\rho}-2\beta f}{\sigma^2(h-f)}-1)}(\rho') = \frac{Z(\frac{2\beta h-2\bar{\rho}}{\sigma^2(h-f)})}{n!} {}_2F_1(-n, n + \frac{2\beta}{\sigma^2} - 1; \frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}; \frac{1 - \frac{2\rho'}{h-f} + \frac{h+f}{h-f}}{2}),$$

where

$$Z(\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}) = (\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)})(\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)} + 1) \cdots (\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)} + n)$$

and ${}_2F_1$ is the hypergeometric function.

Finally, given all the formulae above, we can derive an expression for the probability kernel of the Jacobi process $P(\rho_1, \rho_2; t_1, t_2)$:

$$P(\rho(0), \rho'; 0, \tau) = \sum_{n=0}^{\infty} e^{\lambda_n \tau}(1 - \frac{2\rho'}{h-f} + \frac{h+f}{h-f})^{\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}-1}(1 + \frac{2\rho'}{h-f} - \frac{h+f}{h-f})^{\frac{2\bar{\rho}-2\beta f}{\sigma^2(h-f)}-1}\psi_n(\rho(0))\psi_n(\rho'), \tag{5.4}$$

where $\tau = T - t$. We will use this probability kernel to compute the first three moments of the averaged Jacobi process in the next subsection.

5.2.2 Tailor-type expansion for the price

Let us define a random variable

$$\hat{\rho} = \frac{1}{\tau} \int_0^{\tau} \rho(t) dt.$$

It is possible to show that the price of an option at the time t can be written as

$$C(S_1, S_2, t) = C_{BS}(E\hat{\rho}) + \frac{1}{2} \frac{\partial^2 C_{BS}}{\partial \rho^2}\Big|_{E\hat{\rho}} Var(\hat{\rho}) + \frac{1}{6} \frac{\partial^3 C_{BS}}{\partial \rho^3}\Big|_{E\hat{\rho}} Skew(\hat{\rho}) + \cdots, \tag{5.5}$$

where C_{BS} stands for "Black-Scholes price", i.e. price of an option (that depend on two underlying prices) given the constant correlation coefficient (between these prices). For the spread options, C_{BS} is given by nothing else but the Margrabe formula.

But before we substitute the Margrabe formula to the equation (5.5), we need to compute the moments of the averaged Jacobi process $\hat{\rho}$. Using our expression for the probability kernel of Jacobi process (5.4),

$$E\hat{\rho} = \sum_{n=0}^{\infty} \frac{1 - e^{-\frac{\sigma^2}{2}n(n+\frac{2\beta}{\sigma^2}-1)\tau}}{\frac{\sigma^2}{2}n(n+\frac{2\beta}{\sigma^2}-1)} \times$$

$$\psi_n(\rho(0)) \int_f^h d\rho' (\rho' (1 - \frac{2\rho'}{h-f} + \frac{h+f}{h-f})^{\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}-1} (1 + \frac{2\rho'}{h-f} - \frac{h+f}{h-f})^{\frac{2\bar{\rho}-2\beta f}{\sigma^2(h-f)}-1} \psi_n(\rho')); \quad (5.6)$$

$$E(\hat{\rho}^2) = \frac{2}{\tau^2} \times \sum_{n,m=0}^{\infty} \Bigg(\frac{1 - e^{-\frac{\sigma^2}{2}n(n+\frac{2\beta}{\sigma^2}-1)\tau}}{\frac{\sigma^4}{4}n(n+\frac{2\beta}{\sigma^2}-1)m(m+\frac{2\beta}{\sigma^2}-1)} +$$

$$\frac{e^{-\frac{\sigma^2}{2}n(n+\frac{2\beta}{\sigma^2}-1)\tau} - e^{-\frac{\sigma^2}{2}m(m+\frac{2\beta}{\sigma^2}-1)\tau}}{\frac{\sigma^4}{4}(n(n+\frac{2\beta}{\sigma^2}-1) - m(m+\frac{2\beta}{\sigma^2}-1))m(m+\frac{2\beta}{\sigma^2}-1)} \Bigg) \times$$

$$\psi_n(\rho(0)) \int_f^h \int_f^h d\rho_x d\rho_y \Bigg(\rho_x (1 - \frac{2\rho_x}{h-f} + \frac{h+f}{h-f})^{\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}-1} (1 + \frac{2\rho_x}{h-f} - \frac{h+f}{h-f})^{\frac{2\bar{\rho}-2\beta f}{\sigma^2(h-f)}-1} \times$$

$$\rho_y (1 - \frac{2\rho_y}{h-f} + \frac{h+f}{h-f})^{\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}-1} (1 + \frac{2\rho_y}{h-f} - \frac{h+f}{h-f})^{\frac{2\bar{\rho}-2\beta f}{\sigma^2(h-f)}-1} \psi_n(\rho_x)\psi_m(\rho_x)\psi_m(\rho_y) \Bigg); \quad (5.7)$$

$$E(\hat{\rho}^3) = \frac{6}{\tau^3} \times \sum_{n,m,l=0}^{\infty} \Bigg(\frac{1 - e^{-\frac{\sigma^2}{2} n(n+\frac{2\beta}{\sigma^2}-1)\tau}}{\frac{\sigma^6}{8} n(n+\frac{2\beta}{\sigma^2}-1) m(m+\frac{2\beta}{\sigma^2}-1) l(l+\frac{2\beta}{\sigma^2}-1)} +$$

$$\frac{e^{-\frac{\sigma^2}{2} n(n+\frac{2\beta}{\sigma^2}-1)\tau} - e^{-\frac{\sigma^2}{2} m(m+\frac{2\beta}{\sigma^2}-1)\tau}}{\frac{\sigma^6}{8}(n(n+\frac{2\beta}{\sigma^2}-1) - m(m+\frac{2\beta}{\sigma^2}-1)) m(m+\frac{2\beta}{\sigma^2}-1) l(l+\frac{2\beta}{\sigma^2}-1)} +$$

$$\frac{e^{-\frac{\sigma^2}{2} n(n+\frac{2\beta}{\sigma^2}-1)\tau} - e^{-\frac{\sigma^2}{2} l(l+\frac{2\beta}{\sigma^2}-1)\tau}}{\frac{\sigma^6}{8}(n(n+\frac{2\beta}{\sigma^2}-1) - l(l+\frac{2\beta}{\sigma^2}-1))(m(m+\frac{2\beta}{\sigma^2}-1) - l(l+\frac{2\beta}{\sigma^2}-1)) l(l+\frac{2\beta}{\sigma^2}-1)} +$$

$$\frac{e^{-\frac{\sigma^2}{2} m(m+\frac{2\beta}{\sigma^2}-1)\tau} - e^{-\frac{\sigma^2}{2} n(n+\frac{2\beta}{\sigma^2}-1)\tau}}{\frac{\sigma^6}{8}(n(n+\frac{2\beta}{\sigma^2}-1) - m(m+\frac{2\beta}{\sigma^2}-1))(m(m+\frac{2\beta}{\sigma^2}-1) - l(l+\frac{2\beta}{\sigma^2}-1)) l(l+\frac{2\beta}{\sigma^2}-1)} \Bigg) \times$$

$$\psi_n(\rho(0)) \int_f^h \int_f^h d\rho_x d\rho_y d\rho_z \Bigg(\rho_x (1 - \frac{2\rho_x}{h-f} + \frac{h+f}{h-f})^{\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}-1} (1 + \frac{2\rho_x}{h-f} - \frac{h+f}{h-f})^{\frac{2\bar{\rho} - 2\beta f}{\sigma^2(h-f)}-1} \times$$

$$\rho_y (1 - \frac{2\rho_y}{h-f} + \frac{h+f}{h-f})^{\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}-1} (1 + \frac{2\rho_y}{h-f} - \frac{h+f}{h-f})^{\frac{2\bar{\rho} - 2\beta f}{\sigma^2(h-f)}-1} \times$$

$$\rho_z (1 - \frac{2\rho_z}{h-f} + \frac{h+f}{h-f})^{\frac{2\beta h - 2\bar{\rho}}{\sigma^2(h-f)}-1} (1 + \frac{2\rho_z}{h-f} - \frac{h+f}{h-f})^{\frac{2\bar{\rho} - 2\beta f}{\sigma^2(h-f)}-1} \psi_n(\rho_x)\psi_m(\rho_x)\psi_m(\rho_y)\psi_l(\rho_y)\psi_l(\rho_z) \Bigg). \tag{5.8}$$

These formulae, even though lengthy, allow to approximate the option price almost directly. The infinite sums in the formulae are to be approximated by the finite ones; the integrals from f to h are to be computed numerically. Then one can get the variance and skewness via

$$Var(\hat{\rho}) = E(\hat{\rho}^2) - E(\hat{\rho})^2$$

and

$$Skew(\hat{\rho}) = E(\hat{\rho}^3) - 3E(\hat{\rho}^2)E(\hat{\rho}) + 2E(\hat{\rho})^3.$$

After this is done, one should only keep in mind that the equation (5.5) (taken without $\cdots$) is only an approximation. However, itself it is very accurate, because the forth term in (5.5) is very small - remember that $\hat{\rho}$ is the averaged correlation coefficient, meaning $|\hat{\rho}| \leqslant 1$.

We are now ready to combine the results of the previous sections to get the pricing formula that is, to the best of our knowledge, novel in the literature.

5.2.3 Margrabe formula + the theory above = solution

For the sake of readability, let us remind the Margrabe formula for the price of the spread option with constant drift, volatility and correlation, strike $K = 0$ and maturity T at the time 0:

$$C(S_1, S_2, 0) = S_2(0)\Phi(d_2) - S_1(0)\Phi(d_1), \tag{5.9}$$

where

$$d_1 = \frac{\ln(S_2(0)/S_1(0))}{\sigma_0\sqrt{T}} - \frac{1}{2}\sigma_0\sqrt{T},\ d_2 = \frac{\ln(S_2(0)/S_1(0))}{\sigma_0\sqrt{T}} + \frac{1}{2}\sigma_0\sqrt{T}$$

and

$$\sigma_0 = \sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2}.$$

To compute the derivatives of C, we need

$$\frac{\partial d_2}{\partial \rho} = \frac{\ln(S_2(0)/S_1(0))}{\sqrt{T}}\frac{\partial((\sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2})^{-1})}{\partial \rho} + \frac{1}{2}\frac{\partial(\sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2})}{\partial \rho}\sqrt{T} =$$

$$\frac{\ln(S_2(0)/S_1(0))}{\sqrt{T}}(-\frac{1}{2})\sigma_0^{-3}(-2\sigma_1\sigma_2) + \frac{1}{2}\frac{1}{2}\sigma_0^{-1}(-2\sigma_1\sigma_2)\sqrt{T} = d_1\sigma_0^{-2}\sigma_1\sigma_2.$$

Similarly,

$$\frac{\partial d_1}{\partial \rho} = d_2\sigma_0^{-2}\sigma_1\sigma_2.$$

Using these formulae, we get

$$\frac{\partial^2 d_2}{\partial \rho^2} = \frac{\partial d_1}{\partial \rho}\sigma_0^{-2}\sigma_1\sigma_2 + d_1\frac{\partial(\sigma_0^{-2})}{\partial \rho}\sigma_1\sigma_2 = (2d_1 + d_2)(\sigma_0^{-2}\sigma_1\sigma_2)^2,$$

and

$$\frac{\partial^2 d_1}{\partial \rho^2} = (d_1 + 2d_2)(\sigma_0^{-2}\sigma_1\sigma_2)^2.$$

Given all the above, we can begin to differentiate the Margrabe formula:

$$\frac{\partial(\Phi(d_2))}{\partial \rho} = \frac{1}{\sqrt{2\pi}} e^{-\frac{d_2^2}{2}} \frac{\partial d_2}{\partial \rho},$$

then

$$\frac{\partial^2(\Phi(d_2))}{\partial \rho^2} = \frac{1}{\sqrt{2\pi}}(-d_2 e^{-\frac{d_2^2}{2}}(\frac{\partial d_2}{\partial \rho})^2 + e^{-\frac{d_2^2}{2}} \frac{\partial^2 d_2}{\partial \rho^2}) =$$

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{d_2^2}{2}}(-d_1^2 d_2 + 2d_1 + d_2)(\sigma_0^{-2}\sigma_1\sigma_2)^2 \tag{5.10}$$

and

$$\frac{\partial^3(\Phi(d_2))}{\partial \rho^3} = \frac{1}{\sqrt{2\pi}}(-d_2 e^{-\frac{d_2^2}{2}} \frac{\partial d_2}{\partial \rho}(-d_1^2 d_2 + 2d_1 + d_2)(\sigma_0^{-2}\sigma_1\sigma_2)^2 +$$

$$e^{-\frac{d_2^2}{2}} \frac{\partial(-d_1^2 d_2 + 2d_1 + d_2)}{\partial \rho}(\sigma_0^{-2}\sigma_1\sigma_2)^2 + e^{-\frac{d_2^2}{2}}(-d_1^2 d_2 + 2d_1 + d_2)\frac{\partial(\sigma_0^{-4})}{\partial \rho}(\sigma_1\sigma_2)^2) =$$

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{d_2^2}{2}}(d_1^3 d_2^2 - d_1^3 - 6d_1^2 d_2 - 3d_1 d_2^2 + 9d_1 + 6d_2)(\sigma_0^{-2}\sigma_1\sigma_2)^3 \tag{5.11}$$

One can calculate the derivatives of $\Phi(d_1)$ in the same way to get the symmetric answers:

$$\frac{\partial^2(\Phi(d_1))}{\partial \rho^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}}(-d_1 d_2^2 + d_1 + 2d_2)(\sigma_0^{-2}\sigma_1\sigma_2)^2 \tag{5.12}$$

$$\frac{\partial^3(\Phi(d_1))}{\partial \rho^3} = \frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}}(d_1^2 d_2^3 - d_2^3 - 6d_1 d_2^2 - 3d_1^2 d_2 + 6d_1 + 9d_2)(\sigma_0^{-2}\sigma_1\sigma_2)^3 \tag{5.13}$$

Combining (5.10), (5.11), (5.12) and (5.13),

$$\frac{\partial^2 C}{\partial \rho^2} = \frac{1}{\sqrt{2\pi}}(\sigma_0^{-2}\sigma_1\sigma_2)^2(S_2(0)e^{-\frac{d_2^2}{2}}(-d_1^2 d_2+2d_1+d_2)-S_1(0)e^{-\frac{d_1^2}{2}}(-d_1 d_2^2+d_1+2d_2)) \quad (5.14)$$

and

$$\frac{\partial^3 C}{\partial \rho^3} = \frac{1}{\sqrt{2\pi}}(\sigma_0^{-2}\sigma_1\sigma_2)^3(S_2(0)e^{-\frac{d_2^2}{2}}(d_1^3 d_2^2 - d_1^3 - 6d_1^2 d_2 - 3d_1 d_2^2 + 9d_1 + 6d_2) -$$

$$S_1(0)e^{-\frac{d_1^2}{2}}(d_1^2 d_2^3 - d_2^3 - 6d_1 d_2^2 - 3d_1^2 d_2 + 6d_1 + 9d_2)) \quad (5.15)$$

Now recall that C here is C_{BS} in (5.5). We already know the expectation, variance and skewness in (5.5) from (5.6), (5.7), and (5.8). The last thing to do to get the approximated price is to substitute these, (5.14) and (5.15) to (5.5). We will not do this directly in order to avoid two-pages long final formula. However, let us summarize all the above as a theorem.

Theorem 5.2.1. *Let the prices S_1, S_2 of two assets be given by (5.1), and let C be the price of spread option with strike $K = 0$ and maturity T written on the difference $S_2 - S_1$. Let ρ in (5.1) be a Jacobi process defined in 5.0.1, with the restrictions (5.2) and (5.3). Then C is given by (5.5), where all the values are given by (5.6), (5.7), (5.8), (5.9), (5.14), and (5.15).*

Chapter 6

Conclusion

The world market is affected by an immense amount of different factors, hence the prices change in a hardly predictable way. Mathematical framework provided by stochastic analysis is proven to be an efficient tool to harness this uncertainty. On the other hand, stochastic analysis itself is an area of research for mathematicians. In the first chapters of the thesis, we overviewed its methods from both purely mathematical and financial perspectives.

Application of financial mathematics techniques to derivative pricing leads to a number of interesting models, yet there is always a certain gap between what these models predict and the real-world prices. Therefore, models need to be constantly improved, for a better model means better understanding of the reality. We picked a particular derivative known as spread option. Their prices depend on two other prices that are additionally interdependent. To specify and, in a sense, relax the assumptions we put on this dependance structure seemed like a logical step to improve a model.

The key chapter of the thesis is chapter 5, where we use [26] methods to introduce pricing formula for a spread options written on the difference of two correlated assets, where the correlation itself is driven by a stochastic process. To the best of our knowledge, the formula is novel in the literature, yet there are, of course, more difficulties to overcome. For example, same technique can be used to produce the pricing formula for spread options with stochastic correlation *and volatility*, provided one has a ρ-constant pricing formula to substitute instead of "Black-Scholes price". We leave this application for the further research.

We mostly worked with Wiener process(es). Another interesting direction would be to try to utilize the subordinated (see section 3.2.2) processes instead. An issue here is that, when there are Wiener processes involved, their interdependence structure can be evaluated in the simple correlation. However, imagine a pair of *subordinated* Brownian motions, where the corresponding Wiener processes are correlated. The subordinated processes obviously depend on each other, but more complex dependance measures, e.g. tail dependance, must be used to make sense out of this interconnection.

Last but not least, numerical experiments are needed to validate the models considered in the thesis, to compare them and to figure out when each of them can produce its best results. For stochastic correlation models, an open question is if there are "better" (and in what sense) processes to model the correlation than the Jacobi one.

Bibliography

[1] Ball, C.A., Roma, A., Stochastic Volatility Option Pricing, J. Financial and Quantitative Analysis, 29, 4, 589-607 (1994).

[2] Barndorff-Nielsen, O.E., Normal inverse Gaussian distributions and stochastic volatility modelling, Scandinavian Journal of Statistics, 24, 1-13 (1997).

[3] Benth, F.E., Di Nunno, G., Khedher, A., Schmeck, M.D., Pricing of spread options on a bivariate jump market and stability to model risk, Appl. Math. Finance, 22(1), 28-62 (2015).

[4] Black, F., Scholes, M., The pricing of options and corporate liabilities, J. Polit. Economy, 81, 637-659 (1973).

[5] Carmona, R., Durrleman, V., Pricing and Hedging Spread Options, SIAM Review, 45, 4, 627-685 (2003).

[6] Carr, P., Geman, H., Madan, D.B., Yor, M., The fine structure of asset returns: An empirical investigation, Journal of Business 75, 305-332 (2002).

[7] Cox, J.C., Ross, S.A., The pricing of options for jump processes (working paper), Rodney L. White Center, University of Pennsylvania, Philadelphia 2-75 (1975).

[8] Cox, J.C., Ingersoll, J.E. Jr., Ross, S.A., A theory of the term structure of interest rates, Econometrica, 53, 2, 385-407 (1985).

[9] Delbaen, F., Schachermayer, W., A general version of the fundamental theorem of asset pricing, Math. Ann., 300, 463-520 (1994).

[10] Derman, E., Kani, I., Zou, J.Z., The Local Volatility Surface, Goldman Sachs Quantitative Strategies Research Notes (1995).

[11] Dupire, B., Pricing with a Smile, Risk Volatility, 7, 1 (1994).

[12] Eberlein, E., Papapantoleon, A., Shiryaev, A.N., Esscher transform and the duality principle for multidimensional semimartingales, Ann. Appl. Prob., 19(5), 1944-1971 (2009).

[13] Fischer, S., Call Option Pricing When the Exercise Price is Uncertain, and the Valuation of Index Bonds, J. Finance, 33, 169-176, 1978.

[14] Girsanov, I.V., On transforming a certain class of stochastic processes by absolutely continuous changes of measures, Theory of Probability and Its Applications, 5, 285-301 (1960).

[15] Heston, S., A closed-form solution for options with stochastic volatility with applications to bond and currency options, Review of Financial Studies, 6, 327-343 (1993).

[16] Hull, J.C., White, A., The pricing of options on assets with stochastic volatilities, The Journal of Finance, 42, 281-300 (1987).

[17] Itô, K., Stochastic Integral, Proc. Imp. Acad., 20, Tokyo, 519-524 (1944).

[18] Itô, K., On a formula concerning stochastic differentials, Nagoya Math. J., 3, 55-65 (1951).

[19] Jacod, J., Shiryaev, A. N., (2nd ed.) Limit Theorems for Stochastic Processes, Grundlehren der Mathematischen Wissenschaften (Fundamental Principles of Mathematical Sciences), 288, Springer, Berlin (2003).

[20] Jeanblanc, M., Yor, M., Chesney, M., Mathematical Methods for Financial Markets, Springer (2009).

[21] Kazamaki, N., The equivalence of two conditions on weighted norm inequalities for martingales, In Proc. Intern. Symp. SDE Kyoto 1976 (Itô, K. ed.) Kinokuniya, Tokyo, 141-152 (1978).

[22] Kim, Y.-J., Option pricing under stochastic interest rates: an empirical investigation, Asia-Pacific Financial Markets, 9, 1, 23-44 (2002).

[23] Kreps, D., Arbitrage and equilibrium in economics with infinitely many commodities, Journal of Mathematical Economics 8, 1535 (1981).

[24] Liptser, R.S., Shiryaev, A.N., (Aries A.B., translator, 1977, 1978), (2nd, revised and expanded edition) Statistics of Random Processes, Springer-Verlag, Heidelberg (2001).

[25] Novikov, A.A., On moment inequalities and identities for stochastic integrals, Proceedings of the Second Japan-USSR Symp. Prob. Theory, Lecture Notes in Math. Vol., 330, 333-339 (1973).

[26] Ma, J., Pricing Foreign Equity Options with Stochastic Correlation and Volatility, Ann. Econ. Fin., 10(2), 303-327 (2009).

[27] Madan, D.B., Seneta, E., The variance gamma (VG) model for share market returns, Journal of Business 63, 511-524 (1990).

[28] Margrabe, W., The value of an option to exchange one asset for another, J. Finance, 33, 177-186 (1978).

[29] Markus, L., Mathematics for Pricing Continuously Traded Financial Assets and Their Derivatives (lecture notes), Department of Probability Theory and Statistics, Institute of Mathematics (2016).

[30] R.C. Merton, The theory of rational option pricing, Bell J. Econ. Manag. Sci., 4, 141-183 (1973).

[31] Oksendal, B., (5th edition, corrected printing) Stochastic Differential Equations: An Introduction with Applications, Springer-Verlag, Heidelberg, New York (1998).

[32] Ouwehand P., The Lévy-Itô Decomposition Theorem (translation with notes), arXiv:1506.06624 (2015).

[33] Papapantoleon, A., An introduction to Lévy processes with applications in finance (lecture notes), University of Freiburg (2005).

[34] Poulsen, R., The Margrabe Formula, Encyclopedia of Quantitative Finance (2009).

[35] Scott, L., Option Pricing when the Variance changes randomly, Theory, Estimation, and an Application, J. Financial and Quantitative Analysis, 22, 419-438 (1987).

[36] Shiryaev, A. N., Essentials of stochastic finance: facts, models, theory, World scientific (1999).

[37] Szego G., Jacobi Polynomials Ch.4 in Orthogonal Polynomials, 4th ed. (Providence, RI: American Mathematical Society) (1975).

[38] Tankov, P., Voltchkova, E., Jump diffusion models: a practitioner's guide, Banque et Marchés, 99 (2009).

[39] L. Teng, M. Ehrhardt, M. Gunther, Quanto Pricing in Stochastic Correlation Models, SIAM Conference on Financial Mathematics and Engineering (2016).

[40] Yamada T., Watanabe, Sh., On the uniqueness of the solutions of stochastic differential equations, J. Math. Kyoto Univ., 11, 155-167. (1971).