Eötvös Loránd University

Faculty of Natural Sciences

Elter Adam Gustav

Mathematical models of option pricing

Thesis

Mathematics BSc, applied mathematics specialisation

Supervisor:

Ágnes Backhausz

Department of Probability Theory and Statistics

Budapest, 2025

Contents

Introduction

In this thesis, we look at different ways to price a European call option. Chapter 1 contains the most important definitions and concepts, such as filtrations, (adapted) stochastic processes, and derivative securities in general. In Chapter 2, we consider a discrete-time approach, which allows us to introduce important techniques, such as hedging and risk-free pricing. Chapter 5 is the corresponding continuous-time analogue to the model introduced in Chapter 2; here we derive the Black-Scholes-Merton partial differential equation, a starting point of modern mathematical finance and option pricing. The intermediate chapters contain the mathematical framework that allows us to derive the results in the fifth chapter, namely Brownian motion and Itô calculus. The most unusual result, in our opinion, is that one can integrate stochastic processes with respect to other stochastic processes (Itô integral). Chapter 3 introduces Brownian motion, an essential concept for continuous-time modeling, while Chapter 4 considers the calculus needed, built up from the most basic simple processes to the more general case discussing Itô processes, with the pinnacle being the Itô-Doeblin formula. Finally, Chapter 6 contains an outlook to other derivative securities (exotic options), and ways to price them with other tools than the ones we derived, for example, Monte Carlo simulations, or switching the assumption of the underlying following a geometric Brownian motion to a more market-fitting Lévy process.

Chapter 1

Essentials

This chapter contains the basic mathematical definitions and theorems we used in the later chapters, as well as a few financial concepts we had to include for cohesive purposes.

Definition 1.0.1 *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. We call the sequence $(X_n)_{n\in\mathbb{N}}$ of random variables $X_1, X_2, \ldots$ a discrete-time stochastic process. Now let us fix the positive number T, and define a random variable for every $t \in [0, T]$. We call the collection $\{X(t) \ : \ t \in [0, T]\}$ a continuous-time stochastic process.*

Definition 1.0.2 *[1] Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and let T be a fixed positive number. Suppose that for every $t \in [0, T]$ there is a sigma-algebra $\mathcal{F}(t)$ such that if $s < t$, then $\mathcal{F}(s) \subset \mathcal{F}(t)$; we call $\{\mathcal{F}(t) \ : \ t \in [0, T]\}$ a filtration.*

This definition is for a continuous filtration; one can easily imagine what the discrete-time analogue must be like, and can see an example of it in Definition 1.0.4.

Definition 1.0.3 *[2] Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\mathcal{F} \subset \mathcal{A}$ a sub-sigma algebra and $X : \Omega \to \mathbb{R}$ be a random variable satisfying $\mathbb{E}(|X|) < \infty$. Then, the conditional expectation of X given $\mathcal{F}$, denoted by $\mathbb{E}(X|\mathcal{F})$ is a random variable Y with the properties*
(i) Y is $\mathcal{F}$-measurable
(ii) $\forall \ A \in \mathcal{F} : \ \int_A X \ d\mathbb{P} = \int_A Y \ d\mathbb{P}$
Any Y satisfying (i) and (ii) is said to be a version of $\mathbb{E}(X|\mathcal{F})$.

Definition 1.0.4 *[3] Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration, an increasing sequence of sub-sigma algebras of $\mathcal{A}$. A sequence of integrable random variables $(X_n)_{n\in\mathbb{N}}$ is called a martingale with respect to $(\mathcal{F}_n)$, if for all $n \geq 0$*

(i) X_n *is* $\mathcal{F}_n$*-measurable*
(ii) $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$

Definition 1.0.5 *Let* $X_1, X_2, ...$ *be a sequence of random variables (discrete-time stochastic process), and let* $(\mathcal{F}_n)_{n\in\mathbb{N}}$ *be a filtration. We say that* $(X_n)_{n\in\mathbb{N}}$ *is adapted to this filtration, of for every* $n \in \mathbb{N}$ X_n *is* $\mathcal{F}_n$*-measurable.*

Note: If we don't specify the filtration and just say that (X_n) is an adapted stochastic process, we always think of the natural filtration $\mathcal{F}_n = \sigma\{X_1, ..., X_n\}$, the sigma-algebras generated by the random variables $X_1, ...X_n$.

Definition 1.0.6 *[4] Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space, and* $(X_n)_{n\in\mathbb{N}}$ *be a sequence of random variables and* X *be a random variable. We say that* X_n *converges to* X *in* L^2*, if*

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

Definition 1.0.7 *[5] (Lagrange Mean Value Theorem) Let* $f : \mathbb{R} \to \mathbb{R}$ *be continuous on* $[a, b]$ *and differentiable on* (a, b)*. Then, there exists* $c \in (a, b)$ *such that*

$$f'(c) = \frac{f(b)-f(a)}{b-a}$$

We also want to specify what a *derivative security* is. This is a financial contract between two participants, a buyer and a seller. The value of a derivative security depends on the value of the *underlying asset*, which can be about anything, but in our case, it's a stock. A *European call option* is a derivative security, which conveys to its owner the right, but not the obligation to buy a specified amount of the underlying asset (stock) for a specified price (*strike price)* at a specified time (*maturity date*). The owner pays a *premium* for the option when it's set up. In this text we only consider the European call option, which when exercised, pays the amount $\max(S - K, 0)$, where S is the underlying stock price and K is the strike price. If $K > S$, then the option expires worthless, hence the zero in the defining formula. We summarize this in the following

Definition 1.0.8 *[6] A European call(put) option gives the holder the right, but not the obligation, to buy(sell) an asset at a specified time* t*, for a specified price,* K*. The payout of the option is then* $\max(S - K, 0)$ *(or for a put option* $\max(K - S, 0)$*).*

Lastly, a few financial terms; we do not call these definitions because they are not mathematically precise, and we preserve that notation for mathematical terms.
The *stock market* is where stocks are traded. A stock is a type of security that represents partial ownership in a company. Generally speaking, a stock is considered a risky asset due to its tendency to fluctuate in value. A unit of stock is called a *share*. [6]
The *money market* includes securities that are basically risk-less. The money invested in the money market accumulates *interest rate* over time.
The *interest rate*, denoted by $r \geq 0$, is the reason one would invest in the money market. One dollar invested in the money market will pay off $(1+r)^t$ dollars in t years (if the interest rate is accumulated annually). Similarly, one dollar borrowed from the money market will result in a debt of $(1+r)^t$ in t years. [7]
A *portfolio* is just a collection of securities.
Arbitrage is a trading strategy that starts with no money, has zero probability of losing money, and has a positive probability of making money. Though real-life markets sometimes exhibit arbitrage, it only lasts a short amount of time, because trading takes place to remove it. A key feature of efficient markets is that they don't allow arbitrage. [7]

Chapter 2

Binomial model

The binomial asset-pricing model is interesting to us for two reasons: one is that with sufficient enough time steps it could be used in practice to price derivative securities, and two, because we can develop the arbitrage-free pricing theory and risk-neutral pricing in a relatively easy environment. These concepts form the core of this thesis, and they will be picked up again in Chapter 5, but in a different context.
Our main source while writing this chapter was [7].

2.1 One-period model

In this simple case, we only have one time step and thus two time periods: time zero, which we think of as the present, and time one, the future. At time zero, we know the value of a stock (we denote this by $S(0)$), but at time one we do not; the only thing we know is that it'll be one of two values: $S_1(A)$ or $S_1(B)$, where A and B are the two possible outcomes of a random experiment ([7] thinks of this as a coin toss). Let us denote the probabilities of these events occurring by $p > 0$ and $q = 1 - p > 0$, respectively.

Definition 2.1.1 *We call the positive numbers u and d the up-factor and down-factor, where*

$$u = \frac{S_1(A)}{S(0)}, \quad d = \frac{S_1(B)}{S(0)}$$

We make the assumption $d < u$, otherwise we relabel the fractions defined above.
We only consider the case $S_1(A) \neq S_1(B)$, otherwise the stock movement is not random and the model is not interesting.

A key feature of efficient markets is that wealth cannot be generated without risk, so in short, an efficient market cannot allow arbitrage. This idea lies at the heart of arbitrage-free pricing theory.

Proposition 2.1.2 *[7] In order to rule out arbitrage in the one-period binomial model, we must assume*

$$0 < d < 1 + r < u$$

Proof: $0 < d$ was already assumed in Definition 2.1.1.
Let's assume that $d \geq 1 + r$; then, we can borrow money from the money market and invest it in the stock market by buying stock for $S(0)$. Even in the worst-case scenario for the stock at time one (that is $S_1(B) = dS(0) \geq (1 + r)S(0)$), we can pay back our $(1 + r)S(0)$ debt to the money market, and there is a positive probability of the stock being worth strictly more than our debt (because $uS(0) > dS(0)$), allowing arbitrage.
Now let's assume $u \leq 1 + r$; by "shorting the stock" (that is, we borrow $S(0)$ amounts of stock from the stock market) we can immediately sell our position and invest in the money market. Even in the best-case scenario for the stock at time one we can pay back our $uS(0)$ debt to the stock market and there's a positive probability that our money market account is worth strictly more, again allowing arbitrage. ■
Note: If there is no arbitrage in the one-period model, the inequalities in Definition 2.1.4. must hold true.
Now let us consider a *European call option.* This type of derivative security gives the right, but not the obligation to its owner to buy one share of stock at a preselected time for the *strike price* K. The case most interesting for us is when $S_1(B) < K < S_1(A)$, so at time one the option is worth $\max(S_1 - K, 0)$.
"The fundamental question of option pricing is how much the option is worth at time zero."[7]
The idea behind arbitrage-free pricing is that by trading in the stock and money markets, we replicate the option. This replication process we will call *hedging.* In practice, one can think of hedging as protection against loss of value.
Let's assume we have a derivative security paying $D_1(A)$ or $D_1(B)$ at time one, depending on a random experiment with two possible outcomes A and B. Our goal is to determine a time zero price D_0 for this derivative security.
At time zero, we start with a cash position X_0 and buy Δ_0 shares of stock for $S(0)$ per share, so our position at time zero is

$$X_0 - \Delta_0 S(0)$$

At time one, our position looks a bit different:

$$X_1 = (1+r)(X_0 - \Delta_0 S(0)) + \Delta_0 S_1 = (1+r)X_0 + \Delta_0(S_1 - (1+r)S(0))$$

because our cash position grew by the interest rate and our stock is now worth S_1 per share.

Our goal is to assign a value to X_1 in a way that

$$X_1(A) = D_1(A)$$
$$X_1(B) = D_1(B)$$

Multiplying by $\frac{1}{1+r}$ leaves us with the two equations

$$X_0 + \Delta_0(\tfrac{1}{1+r}S_1(A) - S(0)) = \tfrac{1}{1+r}D_1(A) \quad (1.1)$$
$$X_0 + \Delta_0(\tfrac{1}{1+r}S_1(B) - S(0)) = \tfrac{1}{1+r}D_1(B) \quad (1.2)$$

Now we use a "trick" to solve these two equations, which will turn out to be the basis of risk-neutral pricing. The trick is to multiply the first equation by $\tilde{p}$, the second by $\tilde{q} = 1 - \tilde{p}$ and add them together. Doing so we get

$$X_0 = \Delta_0(\tfrac{1}{1+r}[\tilde{p}S_1(A) + \tilde{q}S_1(B)] - S(0)) = \tfrac{1}{1+r}[\tilde{p}D_1(A) + \tilde{q}D_1(B)]$$

We can make the term multiplying Δ_0 drop out by choosing

$$S(0) = \tfrac{1}{1+r}(\tilde{p}S_1(A) + \tilde{q}S_1(B))$$

and obtain

$$X_0 = \tfrac{1}{1+r}(\tilde{p}D_1(A) + \tilde{q}D_1(B)) \quad (1.3)$$

Solving the equation above for $\tilde{p}$ leaves us with

$$S_0 = \tfrac{1}{1+r}(\tilde{p}uS(0) + \tilde{q}dS(0)) = \tfrac{S(0)}{1+r}(\tilde{p}(u-d) + d)$$
$$\Rightarrow \tilde{p} = \tfrac{1+r-d}{u-d}, \qquad \tilde{q} = \tfrac{u-1-r}{u-d}$$

Now we only need Δ_0; we can simply subtract (1.2) from (1.1) to obtain

$$\Delta_0 = \tfrac{D_1(A) - D_1(B)}{S_1(A) - S_1(B)}$$

We call this the *delta-hedging formula* (see section 5.3).
To summarize, if we start with a cash position X_0 defined by (1.3) and buy Δ_0 shares of stock defined by the delta-hedging formula, then at time one our position's worth will match the amount that D_1 is paying, regardless of how the random experiment turned out, thus we "*hedged a short position in the derivative security.*"[7]
We also conclude that the derivative security should be priced at

$$D_0 = \frac{1}{1+r}(\tilde{p}D_1(A) + \tilde{q}D_1(B))$$

where the numbers $\tilde{p}, \tilde{q}$ are the *risk-neutral probabilities* and the equation itself is the *risk-neutral pricing formula* [7]. By doing so, we don't introduce arbitrage. Any other time zero price will introduce arbitrage to the model.

2.2 Multiperiod model

In this section we generalize the previous section's one-period model to multiple time steps. Every time step a random experiment is conducted, and if the outcome is A, the stock price moves up by the up-factor u; otherwise, it moves down by the down-factor d. Additionally, we have a money market with a constant interest rate r. We assume the inequalities in Proposition 2.1.4 to rule out arbitrage.
The starting stock price we denote by $S(0) > 0$, and the time one stock price by S_1; this could be one of two things: either $S_1(A) = uS(0)$ or $S_1(B) = dS(0)$.
We consider the European call option, a derivative security whose payout at any time n is $\max(S_n - K, 0)$, where K is the strike price. The payout at time n of an arbitrary derivative security we denote by D_n. The portfolio value at time n is denoted by X_n.
The main focus once again is to determine D_0, the time zero price for said derivative security.
In order to do so, let's suppose that we sell the security at time zero for D_0 (this is unknown), and buy Δ_0 shares of stock, leaving us with a position

$$X_0 = D_0 - \Delta_0 S(0)$$

At time one our portfolio is

$$X_1 = \Delta_0 S_1 + (1+r)(D_0 - \Delta_0 S(0))$$

so we have two equations

$$X_1(A) = \Delta_0 S_1(A) + (1+r)(D_0 - \Delta_0 S(0))$$
$$X_1(B) = \Delta_0 S_1(B) + (1+r)(D_0 - \Delta_0 S(0))$$

We want these to be equal to $D_1(A)$ and $D_1(B)$, respectively.
At time one we can decide to buy or sell stock, given how the experiment turned out, so at time two our portfolio becomes

$$X_2 = \Delta_1 S_2 + (1+r)(X_1 - \Delta_1 S_1)$$

leaving us with another four equations

$$X_2(AA) = \Delta_1(A)S_2(AA) + (1+r)(X_1(A) - \Delta_1(A)S_1(A))$$
$$X_2(AB) = \Delta_1(A)S_2(AB) + (1+r)(X_1(A) - \Delta_1(A)S_1(A))$$

$$X_2(BA) = \Delta_1(B)S_2(BA) + (1+r)(X_1(B) - \Delta_1(B)S_1(B))$$
$$X_2(BB) = \Delta_1(B)S_2(BB) + (1+r)(X_1(B) - \Delta_1(B)S_1(B))$$

At the end we have six equations for six unknowns $(D_0, \Delta_0, \Delta_1(A), \Delta_1(B), X_1(A), X_1(B))$
To solve this, we use the trick introduced in the previous chapter: multiply by the risk-neutral probabilities and subtract one equation from the other to obtain

$$X_1(A) = \tfrac{1}{1+r}[\tilde{p}D_2(AB) - \tilde{q}D_2(AA)]$$

so we say that the price of the option if the first experiment's outcome is A is

$$D_1(A) = \tfrac{1}{1+r}[\tilde{p}D_2(AB) - \tilde{q}D_2(AA)]$$

This is another instance of the risk-neutral pricing formula.
The delta-hedging formula becomes

$$\Delta_1(A) = \tfrac{D_2(AB) - D_2(AA)}{S_2(AB) - S_2(AA)}$$

Of course, similar formulas work for the other case, when the experiment's outcome is B.
Finally, we solve the first two equations $X_1(A) = D_1(A)$ and $X_1(B) = D_1(B)$ as seen in the previous chapter, and get the delta-hedging formula for Δ_0 and the risk-neutral pricing formula for D_0.
We summarize this by

Theorem 2.2.1 *[7] Consider an N-period binomial model with $0 < d < 1 + r < u$ and*

$$\tilde{p} = \frac{1+r-d}{u-d}, \qquad \tilde{q} = \frac{u-1-r}{u-d}$$

Let D_N be the value of a derivative security paying off at time N. Define the sequence of random variables $D_{N-1}, D_{N-2}, ..., D_0$ by

$$D_n = \frac{1}{1+r}[\tilde{p}D_{n+1}(A) + \tilde{q}D_{n+1}(B)]$$

and define

$$\Delta_n = \frac{D_{n+1}(A) - D_{n+1}(B)}{S_{n+1}(A) - S_{n+1}(B)}$$

and lastly, define the portfolio values $X_1, X_2, ..., X_N$ by the wealth-equation

$$X_{n+1} = \Delta_n S_{n+1} + (1+r)(X_n - \Delta_n S_n)$$

If we set $X_0 = D_0$, then we will have $X_N = D_N$, regardless of how the random experiments turned out.

Chapter 3

Brownian motion

In this chapter we follow [1], and obtain Brownian motion as the limit of a scaled random walk as the number of steps approaches infinity. To have a better understanding, we start with symmetric random walks, mostly to familiarize ourselves with notation and terminology such as increments, quadratic variation, variance, and the martingale property. Section 2.4. introduces Brownian motion, Section 2.5. considers quadratic variation of the Brownian motion, an important property that distinguishes ordinary calculus from stochastic calculus. Section 2.6. gives an example of how the quadratic variation could be used to estimate the volatility of a Brownian motion-driven asset price.

3.1 Symmetric random walk

Imagine standing on an arbitrarily chosen point of the real plane, with a fair coin in hand; the game is the following: you flip the coin, and if it's heads, you step forward and upwards, otherwise, if it's tails, you step forward and downwards. Informally speaking, this is a symmetric random walk.

Formally, let's assume we have $X_1, X_2, ...$ i.i.d. random variables with $\mathbb{P}(X_j = 1) = p$ and $\mathbb{P}(X_j = -1) = 1 - p$.
We choose $p = \frac{1}{2}$, allowing symmetry.

Definition 3.1.1 *[1] We say $(W_k)_{k\in\mathbb{N}}$ is a symmetric random walk, if*

$$W_k = \sum_{j=1}^{k} X_j$$

Since W_k is the sum of independent random variables, one can think of it as a random variable as well, meaning that $(W_k)_{k \in \mathbb{N}}$ is a discrete-time stochastic process.

Below we included a picture of a python-run simulation of a random symmetric walk with $k = 100$.

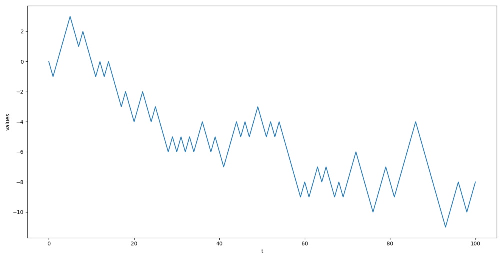

Figure 3.1: Random walk

In order to better understand Brownian motion, we consider some properties of the symmetric random walk, namely *independent increments, martingale property and quadratic variation.*

Definition 3.1.2 *[1] Let $k < l \in \mathbb{N}$; we call the random variable $W_l - W_k = \sum_{j=k+1}^{l} X_j$ the increment of the random symmetric walk between k and l.*

Having independent increments means that if we choose $k_0 < k_1 < ... < k_n, \quad k_i \in \mathbb{N}$, then the random variables

$$W_{k_1} - W_{k_0}, W_{k_2} - W_{k_1}, ..., W_{k_n} - W_{k_{n-1}}$$

are independent. This is due to $W_{k_i} - W_{k_{i-1}} = \sum_{j=k_{i-1}+1}^{k_i} X_j$, so if the intervals are nonoverlapping, the sum depends on different (thus independent) variables.
Moreover, it's easy to see that $\mathbb{E}(W_{k_i} - W_{k_{i-1}}) = 0$, since the expected value is linear and trivially $\mathbb{E}(X_j) = 0 \quad \forall\, j \in \mathbb{N}$.
Similarly, we have $Var(W_{k_i} - W_{k_{i-1}}) = k_i - k_{i-1}$, because $Var(X_j) = \mathbb{E}(X_j^2) = 1$.

We say that *the variance of the symmetric random walk accumulates at rate one per unit time*, meaning that the variance of the increment over a time interval k to l equals $l - k$.

Now let us show the martingale property; let $m < n \in \mathbb{N}$ and suppose $\mathcal{F}_n$ is the corresponding sigma-algebra, the one generated by $\{W_1, W_2, ..., W_n\}$. Then,

$$\mathbb{E}(W_m|\mathcal{F}_n) = \mathbb{E}((W_m - W_n) + W_n|\mathcal{F}_n) =$$
$$= \mathbb{E}(W_m - W_n|\mathcal{F}_n) + \mathbb{E}(W_n|\mathcal{F}_n) =$$
$$= \underbrace{\mathbb{E}(W_m - W_n)}_{0} + W_n = W_n$$

Definition 3.1.3 *[1] Let $n \in \mathbb{N}$; we define the quadratic variation of the random symmetric walk as*

$$[W, W]_n := \sum_{j=1}^{n} (W_j - W_{j-1})^2$$

Note that $[W, W]_n = n$, and unlike the variance, this value is path-dependent. We will see the importance of this quantity when we consider the volatility of Brownian motion-driven asset prices.

3.2 Scaled symmetric random walk

In this section we assume that nt is an integer.
This topic is a simple modification of the previous section's results, but important in the sense that we obtain Brownian motion as the limit of a scaled symmetric random walk as $n \to \infty$. Let's begin with the

Definition 3.2.1 *[1] Let us fix $n \in \mathbb{N}$, and define the scaled random walk as*

$$W^n(t) := \frac{1}{\sqrt{n}} W_{nt}$$

The figure below demonstrates a realization of such a process with $n = 100$ and $t = 5$. One can easily see that in essence we just "speed up the walk" by scaling down the step size and speeding up time.

We define increments and quadratic variation as seen in the previous section; similarly to the random walk, the $W^n(t_{i+1}) - W^n(t_i)$ increments are independent, have zero for expected value and $t_{i+1} - t_i$ for variance. These statements can be verified using the definition, but we'll see it in more detail in the next

Proposition 3.2.2 *[1] The quadratic variation of the scaled symmetric walk up to a time t equals t.*

Proof: Let $t \geq 0$. Then, we have

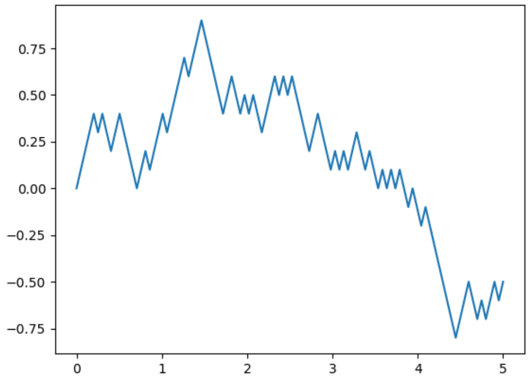

Figure 3.2: $\mathrm{W}^n(t), n = 100, t = 5$

$$[W^{(n)}, W^{(n)}](t) = [\tfrac{1}{\sqrt{n}}W, \tfrac{1}{\sqrt{n}}W](nt) = \sum_{j=1}^{nt}(\tfrac{1}{\sqrt{n}}W_j - \tfrac{1}{\sqrt{n}}W_{j-1})^2 =$$
$$= \sum_{j=1}^{nt}(\tfrac{1}{\sqrt{n}}X_j)^2 = \sum_{j=1}^{nt}\tfrac{1}{n} = t \;\blacksquare$$

Note that although this calculation depends on the path realized, it doesn't really matter which path we take as long as it goes from 0 to t.

3.3 Limiting distribution

As previously mentioned, we're interested in the limit of the scaled random walk as n approaches infinity. For this to make sense, we'll fix the parameter t, write out the moment-generating function of $W^n(t)$, and show that as $n \to \infty$ we approach the moment-generating function of the limiting distribution.
To grasp what a limiting distribution might be, we call upon the Central Limit Theorem, stating that a sequence of independent, identically distributed (i.i.d.) random variables converges in distribution to a normally distributed random variable. We use this fact in the next

Theorem 3.3.1 *[1] Fix $t \geq 0$, let $n \to \infty$. The distribution of $W^n(t)$ converges to the distribution of Z, where Z is a normal distribution with mean zero and variance t.*

Proof: Let us write out the moment-generating function of Z:

$$\varphi_Z(u) = \mathbb{E}(e^{uZ}) = \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty} e^{ux} \cdot e^{\frac{-x^2}{2t}}\, dx$$

Now, since $ux - \frac{x^2}{2t} = -\frac{(x-ut)^2}{2t} + \frac{u^2 t}{2}$, we have

$$\varphi_Z(u) = e^{\frac{u^2 t}{2}} \cdot \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-ut)^2}{2t}} \; dx = e^{\frac{u^2 t}{2}}$$

because the second term in the multiplication is the density function of a normally distributed random variable with mean ut and variance t.

Now let's consider the moment-generating function for $W^{(n)}(t)$:

$$\varphi_{W^{(n)}(t)}(u) = \mathbb{E}(e^{\frac{u}{\sqrt{n}} W_{nt}}) = \mathbb{E}(e^{\frac{u}{\sqrt{n}} \sum_{j=1}^{nt} X_j}) =$$
$$= \mathbb{E}(\prod_{j=1}^{nt} e^{\frac{u}{\sqrt{n}} X_j}) = \prod_{j=1}^{nt} \mathbb{E}(e^{\frac{u}{\sqrt{n}} X_j}) =$$
$$= (\frac{e^{\frac{u}{\sqrt{n}}} + e^{-\frac{u}{\sqrt{n}}}}{2})^{nt}$$

It's easier to show the convergence by taking the logarithms, so we need to show that

$$\lim_{n\to\infty} \; nt \log \frac{e^{\frac{u}{\sqrt{n}}} + e^{-\frac{u}{\sqrt{n}}}}{2} = \frac{u^2 t}{2}.$$

Now let us substitute $x := \frac{1}{\sqrt{n}}$ and write

$$\lim_{x\to 0} \; t \cdot \frac{\log \frac{e^{ux} + e^{-ux}}{2}}{x^2}$$

Both the numerator and the denominator converges to 0, so we can apply L'Hospitals rule:

$$\partial_x \; \log \frac{e^{ux} + e^{-ux}}{2} = \frac{u(e^{ux} - e^{uux})}{e^{ux} + e^{-ux}}$$
$$\partial_x \; x^2 = 2x$$

Thus, we have

$$\lim_{x\to 0} \; \frac{u(e^{ux} - e^{-ux})}{2x(e^{ux} + e^{-ux})}$$

Once again, we need to apply L'Hospitals rule

$$\partial_x \; u(e^{ux} - e^{-ux}) = u^2(e^{ux} + e^{-ux})$$
$$\partial_x \; 2x(e^{ux} + e^{-ux}) = 2(e^{ux} + e^{-ux})$$

And finally, we can compute

$$\lim_{x\to 0} \; t \cdot \frac{\log \frac{e^{ux} + e^{-ux}}{2}}{x^2} = \lim_{x\to 0} \; t \cdot \frac{u^2(e^{ux} + e^{-ux})}{2(e^{ux} + e^{-ux})} = \frac{u^2 t}{2} \; \blacksquare$$

3.4 Brownian motion

At last, we can finally define Brownian motion.

Definition 3.4.1 *[1] Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space, and for* $\omega \in \Omega$ *suppose that there is a continuous function* $B(t),\ t \geq 0$ *that depends on* ω*; we say that* $B(t),\ t \geq 0$ *is a Brownian motion, if for all* $t_0 < t_1 < ... < t_n$ *the increments*

$$B(t_1) - B(t_0), B(t_2) - B(t_1), ..., B(t_n) - B(t_{n-1})$$

are independent and are normally distributed with

$$\mathbb{E}(B(t_i) - B(t_{i-1})) = 0 \text{ and}$$

$$Var(B(t_i) - B(t_{i-1})) = t_i - t_{i-1}$$

An important property of the Brownian motion is that it's a martingale; for this to make sense, we need to define a filtration for the Brownian motion.

Definition 3.4.2 *[1] Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space on which the* $B(t), t \geq 0$ *Brownian motion is defined. A filtration for* $B(t)$ *is a collection of sigma-algebras* $\mathcal{F}(t)$ *satisfying*

1.(Information accumulates) $\mathcal{F}(s) \subset \mathcal{F}(u) \quad \forall\ 0 \leq s < u$

2. (Adaptivity) $\forall\ t \geq 0$ $B(t)$ *is* $\mathcal{F}(t)$*-measurable, meaning that the information available at time* t *is sufficient to evaluate* $B(t)$

3. (Independence of future increments) $\forall\ 0 \leq s < u$ *the increment* $B(u) - B(s)$ *is independent of* $\mathcal{F}(s)$

Theorem 3.4.3 *[1]* $B(t), t \geq 0$ *is a martingale with the filtration defined above.*

Proof: We just have to show the martingale property; let $0 \leq s < u$ and consider the second property (adaptivity)

$$\mathbb{E}(B(u)|\mathcal{F}(s)) = \mathbb{E}(B(u) - B(s) + B(s)|\mathcal{F}(s)) =$$

$$= \mathbb{E}(B(u) - B(s)|\mathcal{F}(s)) + \mathbb{E}(B(s)|\mathcal{F}(s)) =$$

$$= \mathbb{E}(B(u) - B(s)) + B(s) = B(s)\ \blacksquare$$

3.5 Quadratic variation

Previously we computed quadratic variation for the scaled random walk, and we noticed that up to time T it turned out to be exactly T. We derived this result by taking all steps up to time T, squaring them and adding them up.
Now, with Brownian motion, we face a problem: there is no natural step size for the process, so if we want to compute quadratic variation up to time T, any step size is allowed. Thus, we will simply take a large number n, divide T by n, and call $\frac{T}{n}$ our step size. By doing this, we can compute

$$\sum_{j=1}^{n}(B(\frac{jT}{n}) - B(\frac{(j-1)T}{n}))^2$$

Of course, this considers only a discrete amount of partition points, so naturally, we will take the limit of the expression above as $n \to \infty$, reducing the step size and increasing the number of partition points.
The reason we consider quadratic variation in such detail is because it makes stochastic calculus different from ordinary calculus in the sense that ordinary continuous, bounded, real-valued functions have zero quadratic variation, as shown below, unlike Brownian motion, which, although continuous, is nowhere differentiable with respect to the time variable.

3.5.1 First-order variation

To better understand quadratic (second-order) variation, first we introduce first-order variation. Informally speaking, we want to compute the fluctuation of the function f between 0 and T by adding up the up and down movements in a way that the down movements don't subtract but rather add to our term. For this, we partition the interval $[0, T]$, take the difference between the values taken by our function in the partition points, add them up, and take the limit as the number of partition points approaches infinity, or equivalently, as the maximum step size in the partition approaches zero.

Definition 3.5.1 *[1] Let $f : \mathbb{R} \to \mathbb{R}$, and $\mathcal{P}$ be a partition of the interval $[0, T]$ with partition points $0 = t_0 < t_1 < ... < t_n = T$. We denote the maximum step size of $\mathcal{P}$ with $\|\mathcal{P}\| = \max\{t_i - t_{i-1} : i = 1, ..., n\}$ and define the first-order variation of f between 0 and T as*

$$\lim_{\|\mathcal{P}\|\to 0} \sum_{j=1}^{n} |f(t_j) - f(t_{j-1})|$$

Now, if $f : \mathbb{R} \to \mathbb{R}$ is continuous on $[0, T]$ and differentiable on $(0, T)$, then by Lagrange's Mean Value Theorem we have

$$\tfrac{f(t_j)-f(t_{j-1})}{t_j-t_{j-1}} = f'(c_j), \text{ where } c_j \in (t_{j-1}, t_j), \text{ and thus}$$
$$|f(t_j) - f(t_{j-1)}| = |f'(c_j)| \cdot (t_j - t_{j-1})$$

so if we write out the sum in the definition for first-order variation we get

$$\textstyle\sum_{j=1}^{n} |f'(c_j)|(t_j - t_{j-1})$$

which is a Riemann sum for the integral of $|f'|$, and so by definition

$$\lim_{\|\mathcal{P}\|\to 0} \textstyle\sum_{j=1}^{n} |f'(c_j)|(t_j - t_{j-1}) = \int_0^T |f'(t)|\ dt.$$

3.5.2 Second-order variation

Definition 3.5.2 *Let $f : \mathbb{R} \to \mathbb{R}$. We define the quadratic variation of f from 0 to T as*

$$[f, f](T) := \lim_{\|\mathcal{P}\|\to 0} \textstyle\sum_{j=1}^{n} (f(t_j) - f(t_{j-1}))^2$$

Now, if we assume that f has a continuous derivative, then $\int_0^T |f'(t)|^2\ dt$ is finite, and because of

$$\textstyle\sum_{j=1}^{n} (f(t_j) - f(t_{j-1}))^2 = \sum_{j=1}^{n} |f'(c_j)|^2 (t_j - t_{j-1})^2 \leq \|\mathcal{P}\| \cdot \sum_{j=1}^{n} |f'(c_j)|^2 (t_j - t_{j-1})$$

we have

$$[f, f](T) \leq \lim_{\|\mathcal{P}\|\to 0} (\|\mathcal{P}\| \cdot \textstyle\sum_{j=1}^{n} |f'(c_j)|^2 (t_j - t_{j-1})) =$$
$$= \lim_{\|\mathcal{P}\|\to 0} \|\mathcal{P}\| \cdot \lim_{\|\mathcal{P}\|\to 0} \textstyle\sum_{j=1}^{n} |f'(c_j)|^2 (t_j - t_{j-1}) =$$
$$= \lim_{\|\mathcal{P}\|\to 0} \|\mathcal{P}\| \cdot \int_0^T |f'(t)|^2\ dt = 0,$$

meaning that most functions ordinary calculus offers are uninteresting if we consider second-order variation.

In the case of Brownian motion, we can't accept a calculation like the above, because it depends on the Mean Value Theorem, and that assumes the existence of a derivative on $(0, T)$, but our Brownian motion is nowhere differentiable with respect to the time variable, so we need a different approach. Nonetheless, we have

Theorem 3.5.3 *[1] Let $B(t), t \geq 0$ be a Brownian motion, $T \geq 0$. Then,*

$$[B,B](T) = T$$

almost surely.

To prove this in fashion, we define the *kurtosis* of a normal random variable and compute it in the general case to apply it to the special case, when the variables are the increments of a Brownian motion.

Definition 3.5.4 *[1] Let Z be a normal random variable with mean μ and variance σ^2. Then, the kurtosis of Z is*

$$\frac{\mathbb{E}(Z^4)}{Var(Z)^2}$$

Lemma 3.5.5 *[1] Let Z be as in Definition 2.5.4. Then, Z has kurtosis of* 3.

Proof: We can adjust Z by subtracting μ leading to $Z - \mu$ with expected value zero and variance σ^2. Previously, we computed the moment-generating function for $Z-\mu$, that is $\varphi(u) = \mathbb{E}(e^{u(Z-\mu)}) = e^{\frac{u^2\sigma^2}{2}}$.
Differentiating $\varphi(u)$ with respect to u we obtain

$$\varphi'(u) = \mathbb{E}((Z-\mu)e^{u(Z-\mu)}) = u\sigma^2 e^{\frac{u^2\sigma^2}{2}}$$

By substituting $u = 0$ we get

$$\varphi(0) = \mathbb{E}(Z-\mu) = 0$$

Differentiating for a second time we obtain

$$\varphi"(u) = \mathbb{E}((Z-\mu)^2 e^{u(Z-\mu)}) = (\sigma^2 + u^2\sigma^4)e^{\frac{u^2\sigma^2}{2}}$$

Once again, we can simply substitute $u = 0$ to get the second moment of Z.
Differentiating two more times we get

$$\varphi^{(4)}(u) = \mathbb{E}((Z-\mu)^4 e^{u(Z-\mu)}) = (3\sigma^4 + 6\sigma^6u^2 + \sigma^8u^4)e^{\frac{u^2\sigma^2}{2}}$$

And lastly, substituting $u = 0$ we can safely say that $\mathbb{E}((Z-\mu)^4) = 3\sigma^4$, leading to a kurtosis of $\frac{3\sigma^4}{(\sigma^2)^2} = 3$ ■

Proof (of Theorem 2.5.3) We want to compute

$$\lim_{\|\mathcal{P}\|\to 0} \sum_{j=1}^{n} (B(t_j) - B(t_{j-1}))^2$$

so let us denote $\sum_{j=1}^{n}(B(t_j) - B(t_{j-1}))^2$ by Q_n.

Of course Q_n is a random variable, because it depends on the path that the Brownian motion $B(t)$ takes; we will show that the expected value of Q_n is T, regardless of the path, and that the variance of Q_n converges to zero as the number of partition points approaches infinity, so in essence we prove that Q_n converges in L^2 to T.

First we note that since Q_n is the sum of independent random variables, both the expected value and variance of Q_n is the sum of the expected values and variances of $(B(t_j) - B(t_{j-1}))^2 =: b_j^2$, thus

$$\mathbb{E}(Q_n) = \sum_{j=1}^{n} \mathbb{E}(b_j^2) = \sum_{j=1}^{n} Var(b_j) = \sum_{j=1}^{n}(t_j - t_{j-1}) = T$$

To prove L^2 convergence we need to show that

$$Var(Q_n) = \mathbb{E}((Q_n - \mathbb{E}(Q_n))^2) = \mathbb{E}((Q_n - T)^2) \to 0,\ \|\mathcal{P}\| \to 0.$$

As mentioned above,

$$Var(Q_n) = \sum_{j=1}^{n} Var(b_j^2) = \sum_{j=1}^{n} \mathbb{E}((b_j^2 - (t_j - t_{j-1}))^2) =$$
$$= \sum_{j=1}^{n}[\mathbb{E}(b_j^4) - 2(t_j - t_{j-1})\mathbb{E}(b_j^2) + (t_j - t_{j-1})^2] = \sum_{j=1}^{n} \mathbb{E}(b_j^4) - (t_j - t_{j-1})^2$$

To compute the fourth moment of b_j, we take advantage of the fact that b_j is a normal random variable with mean zero and variance $t_j - t_{j-1}$ (See Definition 2.4.1.), so the kurtosis of b_j is 3, meaning that $\mathbb{E}(b_j^4) = 3Var(b_j)^2 = 3(t_j - t_{j-1})^2$, so we can write

$$Var(Q_n) = \sum_{j=1}^{n} 2(t_j - t_{j-1})^2 \leq 2\|\mathcal{P}\| \cdot \sum_{j=1}^{n}(t_j - t_{j-1}) = 2\|\mathcal{P}\|T$$

So, if $\|\mathcal{P}\| \to 0$, then $Var(Q_n) \to 0$, and we conclude that $\lim_{\|\mathcal{P}\|\to 0} \sum_{j=1}^{n}(B(t_j) - B(t_{j-1}))^2 = \mathbb{E}(Q_n) = T$ ■

In conclusion, we say that "Brownian motion accumulates quadratic variation at rate one per unit time" [1]

3.6 Volatility of geometric Brownian motion

As promised, we show how the quadratic variation can be used to approximate the volatility of Brownian motion-driven asset prices.
First, let us define the "asset-price model used in the Black-Scholes-Merton option-pricing formula" [1].

Definition 3.6.1 *[1] Let α and $\sigma > 0$ be constants, and consider the expression*

$$S(t) = S(0)\exp(\sigma B(t) + (1 - \tfrac{\sigma^2}{2})t).$$

We call $S(t), t \geq 0$ a geometric Brownian motion.

Now let $0 \leq T_1 < T_2$ be given, and suppose we have observed $S(t)$ for $T_1 \leq t \leq T_2$ on the partition $T_1 = t_0 < t_1 < ... < t_n = T_2$.
The so-called *log-returns* over the subintervals $[t_{j-1}, t_j]$ are

$$\log \tfrac{S(t_j)}{S(t_{j-1})} = \sigma(B(t_j) - B(t_{j-1})) + (\alpha - \tfrac{\sigma^2}{2})(t_j - t_{j-1})$$

Thus, the *realized volatility* (the sum of the squares of log-returns) is

$$\textstyle\sum_{j=1}^n (\log \tfrac{S(t_j)}{S(t_{j-1})})^2 = \sigma^2 \sum_{j=1}^n (B(t_j) - B(t_{j-1}))^2 +$$
$$\textstyle+(\alpha - \tfrac{\sigma^2}{2})^2 \sum_{j=1}^n (t_j - t_{j-1})^2 + 2\sigma(\alpha - \tfrac{\sigma^2}{2}) \sum_{j=1}^n (B(t_j) - B(t_{j-1}))(t_j - t_{j-1})$$

If we allow the maximum step size to approach zero, the first term of the right-hand side is simply σ^2 times the quadratic variation of the Brownian motion between times T_1 and T_2, that is $T_2 - T_1$.
The second term is $(\alpha - \frac{\sigma^2}{2})^2$ times the quadratic variation of t, which is zero, so that term becomes zero.
Lastly, we have to deal with $\lim_{\|\mathcal{P}\| \to 0} \sum_{j=1}^n (B(t_j) - B(t_{j-1}))(t_j - t_{j-1})$. First, notice that

$$|(B(t_j) - B(t_{j-1}))(t_j - t_{j-1})| \leq \max_{1 \leq j \leq n} |(B(t_j) - B(t_{j-1}))| \cdot (t_j - t_{j-1})$$

so we conclude that

$$\textstyle| \sum_{j=1}^n (B(t_j) - B(t_{j-1}))(t_j - t_{j-1})| \leq \sum_{j=1}^n |(B(t_j) - B(t_{j-1}))(t_j - t_{j-1})| \leq$$
$$\textstyle\leq \max\limits_{1 \leq j \leq n} |(B(t_j) - B(t_{j-1}))| \cdot \sum_{j=1}^n (t_j - t_{j-1}) = \max\limits_{1 \leq j \leq n} |(B(t_j) - B(t_{j-1}))| \cdot (T_2 - T_1)$$

Now, since $B(t)$ is continuous, $\lim_{\|\mathcal{P}\| \to 0} \max\limits_{1 \leq j \leq n} |(B(t_j) - B(t_{j-1}))| = 0$, and thus the third term also becomes zero.
To summarize, if the step size is small enough, the last two terms can be ignored, and the realized volatility is approximately equal to $(T_2 - T_1)\sigma^2$, so we write

$$\textstyle\frac{1}{T_2 - T_2} \sum_{j=1}^n (\log \frac{S(t_j)}{S(t_{j-1})})^2 \approx \sigma^2.$$

Chapter 4

Stochastic calculus

This topic is of great interest to us, because it's stochastic calculus that lies at the heart of sophisticated financial mathematics. Although the subject is important, discussing it in detail is beyond the limits of this thesis, so we direct the reader to [1] for a more thorough discussion. We used mostly [1] to write this chapter.

The goal is to introduce the Itô integral and the Itô-Doeblin formula, define a more general stochastic process than the Brownian motion, namely the Itô process, and use it to build an integral where we can integrate with respect to this stochastic process.

4.1 Itô integral

Let us fix $T > 0$, consider a Brownian motion $B(t), t \geq 0$ and have $\Delta(t), t \geq 0$ as a stochastic process adapted to the sigma-algebra defining $B(t)$. The goal is to try and make sense of what an expression

$$\int_0^T \Delta(t)\ dB(t)$$

could mean. First, notice that if $g(t)$ is a function differentiable with respect to t, then

$$\int_0^T \Delta(t)\ dg(t) = \int_0^T \Delta(t) g'(t)\ dt$$

but in our case, $B(t)$ is not differentiable with respect to the time variable, so we need a different approach.

Doing so, first we define the integral for "simple processes" [1], and extend it as a limit for more general functions.

Definition 4.1.1 *[1] We say $\Delta(t)$ is a simple process, if $\Delta(t)$ is constant for every $t \in [t_j, t_{j+1})$, where $0 = t_0 < t_1 < ... < t_n = T$ is a partition of $[0, T]$.*

It might not be obvious at first glance, but $\Delta(t)$ depends on the path that the Brownian motion takes, so it is, in nature, random (except for $\Delta(0)$, because $\Delta(t)$ depends on the information available at time t, and there is no information available at time zero).
We think of the t_j -s as trading times, and the relationship between $B(t)$ and $\Delta(t)$ as the relationship between a stock price and the number of shares taken in the stock.
Thus, the gain from trading at different times is

$$I(t) = \Delta(t_0)[B(t) - B(t_0)], \quad 0 \leq t \leq t_1$$

$$I(t) = \Delta(t_0)B(t_1) + \Delta(t)[B(t) - B(t_2)], \quad t_1 \leq t \leq t_2$$

It's easy to see the pattern emerging from this, so we conclude

Definition 4.1.2 *[1] The Itô integral of the simple process $\Delta(t)$ is*

$$I(t) = \int_0^T \Delta(u)\ dB(u) = \sum_{j=1}^n \Delta(t_{j-1})[B(t_j) - B(t_{j-1})] + \Delta(t_n)[B(t) - B(t_{n-1})]$$

One can think of $I(t)$ as a stochastic process in t (the upper limit of integration). Doing so, we obtain interesting results, such as

Lemma 4.1.3 *[1] $I(t), t \geq 0$ is a martingale.*

Because of $I(0) = 0$ and the martingale property, we have $\mathbb{E}(I(t)) = 0\ \forall\ t \geq 0$, and $Var(I(t))$ can be evaluated by the Itô-isometry:

Theorem 4.1.4 *[1] Let $\Delta(t), t \geq 0$ be a simple process and $I(t)$ the Itô integral for the simple process. Then,*

$$\mathbb{E}(I^2(t)) = \mathbb{E}(\int_0^t \Delta(u)\ dB(u)).$$

We have shown in the previous chapter that "Brownian motion accumulates quadratic variation at rate one per unit time" [1]. Now, $B(t)$ is scaled down by $\Delta(t)$ as it enters $I(t) = \int_0^t \Delta(u)dB(u)$, so it's not that surprising that since the increments are squared in the computation of quadratic variation, the quadratic variation of $I(t)$ will depend somehow on $\Delta^2(u)$.

Theorem 4.1.5 *[1] The quadratic variation accumulated up to time t by the Itô integral is*

$$[I, I](t) = \int_0^t \Delta^2(u)\ du$$

Proof: The idea is that first we look at the subintervals $[t_{j-1}, t_j)$, where $\Delta(u)$ is constant, compute the quadratic variation, and then add the terms up.
To do this, we take the partition $t_{j-1} = s_0 < s_1 < ... < s_m = t_j$, and consider

$$\sum_{k=1}^m (I(s_k) - I(s_{k-1}))^2 = \sum_{k=1}^m (\Delta(t_{k-1})(B(t_k) - B(t_{k-1})))^2 =$$
$$= \Delta^2(t_{k-1}) \sum_{k=1}^m (B(t_k) - B(t_{k-1}))^2$$

If we allow m to approach infinity, the maximal distance between partition points approaches zero, so the sum converges to the quadratic variation accumulated by the Brownian motion $B(t)$ between t_{k-1} and t_k, which of course is $t_{k-1} - t_k$.
Thus, we can write

$$\sum_{k=1}^m (I(s_k) - I(s_{k-1}))^2 = \Delta^2(t_{k-1})(t_{k-1} - t_k) = \int_{t_{k-1}}^{t_k} \Delta^2(u)\ du$$

because $\Delta(u) = \Delta(t_{k-1})\ \ \forall\ u \in [t_{k-1}, t_k)$.
This holds true for all subintervals, so we conclude that

$$[I, I](t) = \int_{t_0}^{t_1} \Delta^2(u)\ du + \int_{t_1}^{t_2} \Delta^2(u)\ du + ... = \int_0^t \Delta^2(u)\ du \quad \blacksquare$$

Now we consider more general functions as integrands, namely $\Delta(t)$ is allowed to be continuous and have jumps. We only assume that it is adapted to the same filtration as the simple process and the "square-integrability condition"[1] $\mathbb{E}[\int_0^T \Delta^2(t)\ dt] < \infty$.
As promised, we expand the Itô integral to more general integrands by taking a limit; in order to do so we first approximate $\Delta(t)$ by a series made of simple processes $\Delta_n(t)$. The construction is the following: partition $[0, T]$ by the partition points $0 = t_0 < t_1 < ... < t_n = T$, and let $\Delta_n(t) = \Delta(t_j)$ for $t_j < t < t_{j+1}\ \ \forall\ j \in \{0, 1, ..., n-1\}$. As the number of partition points approaches infinity, we get a better and better approximation for $\Delta(t)$ in the sense that

$$\lim_{n\to\infty} \mathbb{E}(\int_0^T |\Delta_n(t) - \Delta(t)|^2\ dt) = 0$$

Since the integral has been defined for simple processes, we simply define it for $\Delta(t)$ as the limit

$$\int_0^T \Delta(u)\ dB(u) = \lim_{n\to\infty} \int_0^T \Delta_n(u)\ dB(u)$$

We think of the integral once again as a stochastic process in its upper limit of integration $0 \le t \le T$. This allows us to derive the following results

Theorem 4.1.6 *[1] Let T be a fixed positive constant, $\Delta(t), t \geq 0$ an adapted stochastic process that satisfies the square-integrability condition, and let $I(t) = \int_0^t \Delta(u)\ dB(u)$, where the integral is the limit defined above. Then, $I(t)$, thought of as a stochastic process in t has the following properties:*
(i) $I(t)$ is continuous
(ii) $I(t)$ is adapted to the filtration $\mathcal{F}(t)\ \forall\ t \geq 0$
(iii) $I(t)$ is linear
(iv) $I(t)$ is a martingale
(v) $\mathbb{E}(I^2(t)) = \mathbb{E}(\int_0^t \Delta^2(u)\ du)$
(vi) $[I, I](t) = \int_0^t \Delta^2(u)\ du$

Proposition 4.1.7 *[1] Let $B(t), t \geq 0$ be a Brownian motion, $T > 0$ fixed and $0 \leq t \leq T$. Then*

$$\int_0^t B(u)\ dB(u) = \tfrac{1}{2}B^2(t) - \tfrac{1}{2}t$$

One way to prove this is to use the definition above, construct a series of simple processes approximating the Brownian motion and evaluate the limit of the integrals. We will follow a different path and use the Itô-Doeblin formula (see next section), but the details can be found in [1].

4.2 Itô-Doeblin formula

Previously we learned to integrate with respect to Brownian motion. The goal now is to differentiate Brownian motion somehow; formally speaking, we want to make sense of the expression

$$\tfrac{d}{dt}f(B(t))$$

where f is differentiable, and $B(t), t \geq 0$ is a Brownian motion, as usual.
Note that if $B(t)$ was differentiable with respect to t (which is not the case), then the chain rule from real analysis would give us

$$\tfrac{d}{dt}f(B(t)) = f'(B(t)) \cdot B'(t)$$

or in differential form

$$df(B(t)) = f'(B(t)) \cdot B'(t)\ dt = f'(B(t))\ dB(t)$$

But because $B(t), t \geq 0$ has nonzero quadratic variation, there's an additional term entering the differential form:

$$df(B(t)) = f'(B(t))\ dB(t) + \frac{1}{2}f"(B(t))\ dt$$

This we call the *Itô-Doeblin formula in differential form.* [1]
Integrating both sides gives us the *Itô-Doeblin formula in integral form*, concretely

$$f(B(t)) - f(B(0)) = \int_0^t f'(B(u))\ dB(u) + \frac{1}{2}\int_0^t f''(B(u))\ du$$

where the right-hand side consists of an Itô integral (see Theorem 3.1.6.) and an ordinary Lebesgue integral, with respect to the time variable. We generalize this to functions allowing both t and x in their domain.

Theorem 4.2.1 *[1] (Itô-Doeblin formula for Brownian motion) Let $f(t,x)$ be a function for which the partial derivatives $\partial_t f, \partial_x f, \partial_x^2 f$ exist and are continuous, and let $B(t), t \geq 0$ be a Brownian motion. Then, for every $T \geq 0$ we have*

$$f(T, B(T)) = f(0, B(0)) + \int_0^T \partial_t f(t, B(t))\ dt + \int_0^T \partial_x f(t, B(t))\ dB(t) + \frac{1}{2}\int_0^T \partial_x^2 f(t, B(t))\ dt$$

First, we show how the formula works for the nice quadratic function $f(x,t) = \frac{x^2}{2}$, so f doesn't actually depend on t.
Let x_j, x_{j-1} be arbitrarily chosen points; then, using Taylor's expansion

$$f(x_j) - f(x_{j-1}) = f'(x_{j-1})(x_j - x_{j-1}) + \frac{1}{2}f''(x_{j-1})(x_j - x_{j-1})^2$$

Of course, in our special case $f'(x) = x$ and $f''(x) = 1$, so the higher derivatives are all equal to zero, meaning that this expansion is exact.
Now let's fix $T > 0$ and consider the partition $\mathcal{P} = \{t_0, t_1, ..., t_n\}$ of the interval $[0, T]$. We want to compute the change in $f(B(t))$ between times 0 and T. This can be expressed by summing the changes $f(B(t_j)) - f(B(t_{j-1}))$ over all subintervals:

$$f(B(T)) - f(B(0)) = \sum_{j=1}^n (f(B(t_j)) - f(B(t_{j-1})))$$

and by Taylor's formula

$$\sum_{j=1}^n (f(B(t_j)) - f(B(t_{j-1}))) =$$
$$= \sum_{j=1}^n f'(B(t_{j-1}))(B(t_j) - B(t_{j-1})) + \frac{1}{2}\sum_{j=1}^n f''(B(t_{j-1}))(B(t_j) - B(t_{j-1}))^2$$

In our case, this boils down to

$$\sum_{j=1}^{n} B(t_{j-1})(B(t_j) - B(t_{j-1})) + \frac{1}{2}\sum_{j=1}^{n}(B(t_j) - B(t_{j-1}))^2$$

By letting $\|\mathcal{P}\| \to 0$ the term $f(B(T)) - f(B(0))$ is unaffected, while

$$\lim_{\|\mathcal{P}\|\to 0} \sum_{j=1}^{n} B(t_{j-1})(B(t_j) - B(t_{j-1})) = \int_0^T B(t)\ dB(t) \text{ and}$$
$$\lim_{\|\mathcal{P}\|\to 0} \sum_{j=1}^{n}(B(t_j) - B(t_{j-1}))^2 = [B, B](T) = T$$

so we have

$$f(B(T)) - f(B(0)) = \int_0^T B(t)\ dB(t) + \frac{1}{2}T =$$
$$= \int_0^T f'(B(t))\ dB(t) + \frac{1}{2}\int_0^T f''(B(t))\ dt$$

which is the Itô-Doeblin formula for $f(x) = \frac{x^2}{2}$.
For the more general case $f(t, x)$ Taylor's formula states that

$$f(t_j, x_j) = f(t_{j-1}, x_{j-1}) + \partial_t f(t_{j-1}, x_{j-1})(t_j - t_{j-1}) + \partial_x f(t_{j-1}, x_{j-1})(x_j - x_j)+$$
$$+\partial_t\partial_x f(t_{j-1}, x_{j-1})(x_j - x_{j-1})(t_j - t_{j-1})+$$
$$+\frac{1}{2}\partial_x^2 f(t_{j-1}, x_{j-1})(x_j - x_{j-1})^2 + \frac{1}{2}\partial_t^2 f(t_{j-1}, x_{j-1})(t_j - t_{j-1})^2 + \text{higher-order terms}$$

The higher order terms all contain $(x_j - x_{j-1})$ or $(t_j - t_{j-1})$ on at least the second power (thus, later when we substitute $x_j = B(t_j)$ and take limit as $\|\mathcal{P}\| \to 0$, these terms converge to zero, because one can always take out terms such as $\max(t_j - t_{j-1})$).
This gives us (by substituting $x_j = B(t_j)$ and taking the sum over all subintervals)

$$f(T, B(T)) - f(0, B(0)) = \sum_{j=1}^{n}(f(t_j, B(t_j)) - f(t_{j-1}, B(t_{j-1}))) =$$
$$= \sum_{j=1}^{n} \partial_t f(t_{j-1}, B(t_{j-1}))(t_j - t_{j-1})+$$
$$+\sum_{j=1}^{n} \partial_x f(t_{j-1}, B(t_{j-1}))(B(t_j) - B(t_{j-1}))+$$
$$+\sum_{j=1}^{n} \partial_t\partial_x f(t_{j-1}, B(t_{j-1}))(B(t_j) - B(t_{j-1}))(t_j - t_{j-1})+$$
$$+\frac{1}{2}\sum_{j=1}^{n} \partial_x^2 f(t_{j-1}, B(t_{j-1}))(B(t_j) - B(t_{j-1}))^2+$$
$$+\frac{1}{2}\sum_{j=1}^{n} \partial_t^2 f(t_{j-1}, B(t_{j-1}))(t_j - t_{j-1})^2 + \text{higher-order terms}$$

By allowing $\|\mathcal{P}\| \to 0$, we get

$$\lim_{\|\mathcal{P}\|\to 0} \sum_{j=1}^{n} \partial_t f(t_{j-1}, B(t_{j-1}))(t_j - t_{j-1}) = \int_0^T \partial_t f(t, B(t))\ dt$$

so the first term just adds an ordinary Lebesgue integral.

$$\lim_{\|\mathcal{P}\|\to 0}\sum_{j=1}^{n}\partial_x f(t_{j-1},B(t_{j-1}))(B(t_j)-B(t_{j-1}))=\int_0^T \partial_x f(t,B(t))\ dB(T)$$

meaning that the second term gives us the Itô integral in the formula.

$$\lim_{\|\mathcal{P}\|\to 0}|\sum_{j=1}^{n}\partial_t\partial_x f(t_{j-1},B(t_{j-1}))(B(t_j)-B(t_{j-1}))(t_j-t_{j-1})|\leq$$
$$\leq \lim_{\|\mathcal{P}\|\to 0}\max_{1\leq k\leq n}|B(t_k)-B(t_{k-1})|\cdot\lim_{\|\mathcal{P}\|\to 0}\sum_{j=1}^{n}|\partial_t\partial_x f(t_{j-1},B(t_{j-1}))(t_j-t_{j-1})|=$$
$$=0\cdot\int_0^T|\partial_t\partial_x f(t,B(t))|\ dt=0$$

because $B(t)$ is continuous.

$$\lim_{\|\mathcal{P}\|\to 0}\sum_{j=1}^{n}\partial_x^2 f(t_{j-1},B(t_{j-1}))(B(t_j)-B(t_{j-1}))^2=\int_0^T \partial_x^2 f(t,B(t))\ dt$$

because

$$\lim_{\|\mathcal{P}\|\to 0}(B(t_j)-B(t_{j-1}))^2=t_j-t_{j-1}$$

so we can substitute t_j-t_{j-1} and the fourth term gives us the second Lebesgue integral in the formula. Note that although this is not an exact substitution, it gives the same limit. Finally

$$\lim_{\|\mathcal{P}\|\to 0}|\sum_{j=1}^{n}\partial_t^2 f(t_{j-1},B(t_{j-1}))(t_j-t_{j-1})^2|=0$$

and the higher-order terms also contribute zero.

Though this is not a precise mathematical proof, it gives us an understanding of *how and why* the formula works.
With this tool in hand we can give an elegant proof for Proposition 4.1.7. Let $f(x)=\frac{x^2}{2}$ and consider $f(B(t))$. According to the Itô-Doeblin formula

$$f(B(t))=\tfrac{1}{2}B^2(t)=\tfrac{1}{2}B^2(0))+\int_0^t f'(B(u))\ dB(u)+\tfrac{1}{2}\int_0^t f''(B(u))\ du=$$
$$=0+\int_0^t B(u)\ dB(u)+\tfrac{1}{2}\int_0^t B(u)\ du=$$
$$=\int_0^t B(u)\ dB(u)+\tfrac{1}{2}t$$

Rearranging the sides we acquire $\int_0^t B(u)\ dB(u)=\frac{1}{2}B^2(t)-\frac{1}{2}t$, as Proposition 4.1.7. states.

4.3 Itô processes

We want to expand our integral definition for more general stochastic processes than the Brownian motion. The goal of this section is to introduce this more general process, the Itô process, and then define the Itô integral with respect to Itô processes and give an even more general form of the Itô-Doeblin formula.
We begin with the

Definition 4.3.1 *[1] Let $B(t), t \geq 0$ be a Brownian motion and $\mathcal{F}(t), t \geq 0$ be the associated filtration. An Itô process is a stochastic process of the form*

$$X(t) = X(0) + \int_0^t \Delta(u)\ dB(u) + \int_0^t \Gamma(u)\ du$$

where $X(0)$ is a nonrandom constant, $\Delta(u)$ and $\Gamma(u)$ are adapted stochastic processes satisfying

$$\mathbb{E}(\int_0^t \Delta^2(u)\ du) < \infty \text{ and } \int_0^t |\Gamma(u)|\ du < \infty \quad \forall\ t > 0$$

As expected, first we compute the quadratic variation of the Itô process to understand the volatility associated with it.

Lemma 4.3.2 *[1] The quadratic variation of the Itô-process $X(t), t \geq 0$ is*

$$[X, X](t) = \int_0^t \Delta^2(u)\ du$$

Proof: Let us denote $\int_0^t \Delta(u)\ dB(u)$ by $I(t)$ and $\int_0^t \Gamma(u)\ du$ by $J(t)$. Thus, we can write $X(t) = I(t) + J(t)$. Both these processes are continuous in their upper limit of integration, namely t.
Now consider the partition $\mathcal{P} = \{t_0, t_1, ..., t_n\}$ of the interval $[0, t]$.
Naturally

$$\sum_{j=1}^n (X(t_j) - X(t_{j-1}))^2 =$$

$$= \sum_{j=1}^n (I(t_j) - I(t_{j-1}))^2 + 2 \cdot \sum_{j=1}^n (I(t_j) - I(t_{j-1}))(J(t_j) - J(t_{j-1})) + \sum_{j=1}^n (J(t_j) - J(t_{j-1}))^2$$

By allowing $\|\mathcal{P}\| \to 0$ the first term becomes

$$\lim_{\|\mathcal{P}\| \to 0} \sum_{j=1}^n (I(t_j) - I(t_{j-1}))^2 = [I, I](t) = \int_0^t \Delta^2(u)\ du$$

so now we just have to show that the remaining terms contribute zero, and the proof is complete. First,

$$\lim_{\|\mathcal{P}\|\to 0} |\sum_{j=1}^{n}(J(t_j) - J(t_{j-1}))^2| \leq$$
$$\leq \lim_{\|\mathcal{P}\|\to 0} \max_{1\leq k\leq n} |J(t_k) - J(t_{k-1})| \cdot \lim_{\|\mathcal{P}\|\to 0} \sum_{j=1}^{n} |J(t_j) - J(t_{j-1})| \leq$$
$$\leq 0 \cdot \int_0^t |\Gamma(u)|\ du = 0$$

because $J(u)$ is continuous and the integral is finite.
Similarly

$$\lim_{\|\mathcal{P}\|\to 0} |2 \cdot \sum_{j=1}^{n}(I(t_j) - I(t_{j-1}))(J(t_j) - J(t_{j-1}))| \leq$$
$$\leq \lim_{\|\mathcal{P}\|\to 0} \max_{1\leq k\leq n} |I(t_k) - I(t_{k-1})| \cdot 2 \lim_{\|\mathcal{P}\|\to 0} \sum_{j=1}^{n} |J(t_j) - J(t_{j-1})| \leq$$
$$\leq 0 \cdot \int_0^t |\Gamma(u)|\ du = 0$$

because $I(t)$ is also continuous and the integral is finite. ■
Writing out the equation defining the Itô process in differential form we obtain

$$dX(t) = \Delta(t)\ dB(t) + \Gamma(t)\ dt$$

and taking advantage of

$$dB(T)\ dB(t) = dt, \quad dB(t)\ d(t) = 0, \quad dt\ dt = 0$$

we can write

$$dX(t)\ dX(t) = \Delta^2(t)\ dB(t)\ dB(t) + 2\Delta(t)\Gamma(t)\ dB(t)\ dt = \Gamma^2(t)\ dt\ dt =$$
$$= \Delta^2(t)\ dt$$

summarizing the results of Lemma 4.3.2.
Though the equations above are not precisely defined, intuitively one can understand the meaning behind them: $dB(t)\ dB(t) = dt$ captures the "Brownian motion accumulates quadratic variation at rate one per unit time" [1], $dB(t)\ dt = 0$ means that the variation of the cross-product is zero (see section 3.6.), and so is $dt\ dt$.
The point of writing out the equations in differential form is that it might be easier to memorize them, and to break up $dX(t)$ into two parts when we want to define an integral with respect to an Itô process (see the definition below).

Definition 4.3.3 *[1] Let $X(t), t \geq 0$ be an Itô process as in Definition 4.3.1., and let $\Lambda(t), t \geq 0$ be an adapted stochastic process satisfying*

$$\int_0^t \Lambda^2(u)\Delta^2(u)\ du < \infty \textit{ and } \int_0^t |\Lambda(u)\Gamma(u)|\ du < \infty$$

The integral with respect to this Itô process is

$$\int_0^t \Lambda(u)\ dX(u) = \int_0^t \Lambda(u)\Delta(u)\ dB(u) + \int_0^t \Lambda(u)\Gamma(u)\ du$$

Of course we need the corresponding Itô-Doeblin formula, so consider

Theorem 4.3.4 *[1] (Itô-Doeblin formula for Itô processes) Let $X(t), t \geq 0$ be an Itô-process and let $f(t,x)$ be a differentiable function for which the partial derivatives $\partial_t f, \partial_x f, \partial_x^2 f$ exist and are continuous. Then, for every $T \geq 0$*

$$f(T, X(T)) = = f(0, X(0)) + \int_0^T \partial_t f(t, X(t))\ dt + \int_0^T \partial_x f(t, X(t))\Delta(t)\ dB(t) + \\ + \int_0^T \partial_x f(t, X(t))\Gamma(t)\ dt + \tfrac{1}{2}\int_0^T \partial_x^2 f(t, X(t))\Delta^2(t)\ dt$$

Note: the expression above could be written in the more compressed form

$$f(0, X(0)) + \int_0^T \partial_t(t, X(t))\ dt + \int_0^T \partial_x f(t, X(t))\ dX(t) + \tfrac{1}{2}\int_0^T \partial_x^2 f(t, X(t))\ d[X, X](t)$$

The reasoning behind the formula being correct is the same as in Theorem 4.2.1.
Note: In Chapter 5, we use a slightly different form of the above formula, namely the *differential form* [1]

$$df(t, X(t)) = \partial_t f(t, X(t))\ dt + \partial_x f(t, X(t))\ dX(t) + \tfrac{1}{2}\partial_x^2 f(t, X(t))\ dX(t)\ dX(t)$$

Chapter 5

Black-Scholes-Merton equation

In this chapter we derive the Black-Scholes-Merton partial differential equation for the price of an option, where the underlying asset is modeled by a geometric Brownian motion (see definition 3.6.1). The idea is similar to the one we used in Chapter 2: we want to determine the initial wealth required to hedge a short position in the option. Once again, we follow [1] throughout the entire chapter.

5.1 Evolution of portfolio value

We start by defining our portfolio: we denote the portfolio value at time t by $X(t)$. The portfolio consists of a money market investment paying a constant interest rate r, and a stock market investment, where the stock $S(t)$ is modeled by a geometric Brownian motion, so we write

$$dS(t) = \alpha S(t)\ dt + \sigma S(t)\ dB(t)$$

where α and σ are constants, and $B(t), t \geq 0$ is the Brownian motion.
The number of shares held at time t we denote by $\Delta(t)$; this stochastic process is adapted to the same filtration that defines $B(t)$.
It's important to note that the remainder

$$X(t) - \Delta(t)S(t)$$

is always invested in the money market.
The change in our portfolio value is due to two factors: the change in the stock, and the change in the money market asset (an interest is being paid), so we write

$$dX(t) = \Delta(t)\ dS(t) + r(X(t) - \Delta(t)S(t))\ dt$$

Substituting into $dS(t)$ we obtain

$$dX(t) = \Delta(t)(\alpha S(t)\ dt + \sigma S(t)\ dB(t)) + r(X(t) - \Delta(t)S(t))\ dt =$$
$$= rX(t)\ dt + \Delta(t)(\alpha - r)S(t)\ dt + \Delta(t)\sigma S(t)\ dB(t)$$

This equation could be interpreted the following way: the change in the portfolio is first the average rate of return on the portfolio $rX(t)\ dt$, second the risk premium for investing in the stock (second term), and third the volatility proportional to the size of the investment (third term). [1]
The analogue to this equation is the wealth equation seen in Chapter 2.
Let us introduce the discounted stock price

Definition 5.1.1 *[1] We call* $e^{-rt}S(t)$ *the discounted stock price.*

Similarly,

Definition 5.1.2 *[1] We call* $e^{-rt}X(t)$ *the discounted portfolio value.*

Let $f(t, x) = e^{-rt}$; considering the Itô-Doeblin formula from the previous chapter we have

$$d(e^{-rt}S(t)) = df(t, S(t)) =$$
$$= \partial_t f(t, S(t))\ dt + \partial_x f(t, S(t))\ dS(t) + \tfrac{1}{2}\partial_x^2 f(t, S(t))\ dS(t)\ dS(t) =$$
$$= -re^{-rt}S(t)\ dt + e^{-rt}\ dS(t) =$$
$$= -re^{-rt}S(t)\ dt + e^{-rt}[\alpha S(t)\ dt + \sigma S(t)\ dB(t)] =$$
$$= (\alpha - r)e^{-rt}S(t)\ dt + \sigma e^{-rt}S(t)\ dB(t)$$

This we will call the *differential of the discounted stock price.* [1]
The differential for the discounted portfolio value is

$$d(e^{-rt}X(t)) = df(t, X(t)) =$$
$$= \partial_t f(t, X(t))\ dt + \partial_x f(t, X(t))\ dX(t) + \tfrac{1}{2}\partial_x^2 f(t, X(t))\ dX(t)\ dX(t) =$$
$$= -re^{-rt}X(t)\ dt + e^{-rt}\ dX(t) =$$
$$= -re^{-rt}X(t)\ dt + e^{-rt}[rX(t)\ dt + \Delta(t)(\alpha - r)S(t)\ dt + \Delta(t)\sigma S(t)\ dB(t)] =$$
$$= e^{-rt}\Delta(t)(\alpha - r)S(t)\ dt + e^{-rt}\sigma\Delta(t)S(t)\ dB(t) =$$
$$= \Delta(t)\ d(e^{-rt}S(t))$$

So in essence discounting reduces the mean rate of return form α to $\alpha - r$ in the stock price, and completely removes the underlying rate of return r from the portfolio value; $\Delta(t)\ d(e^{-rt}S(t))$ shows us that the change in the discounted portfolio value is only because of the change in the discounted stock price. [1]

5.2 Evolution of option value

We only consider a European call option with payout $\max(S(T) - K, 0)$, where T is a prefixed time (maturity date/expiration date), and $K > 0$ is a prefixed constant (strike price). The value of such an option only depends on a few factors: the time left to maturity (when the option can be exercised), the value of the underlying asset at that time, and the parameters r, σ, K. Out of the above, only time and stock value vary after setting up the contract defining the option.
So, let us consider a function $c(t, x)$; this will determine the value of our option at time t if the stock price at that time is $x = S(t)$. Substituting $S(t)$ we have $c(t, S(t))$, a continuous-time stochastic process.
Our main goal is to determine $c(t, x)$, so that we have a formula for future option values in terms of future stock prices. [1]
First, let us compute the differential of $c(t, S(t))$; the Itô-Doeblin formula gives us

$$dc(t, S(t)) =$$
$$= \partial_t c(t, S(t))\ dt + \partial_x c(t, S(t))\ dS(t) + \tfrac{1}{2}\partial_x^2 c(t, S(t))\ dS(t)\ dS(t) =$$
$$= \partial_t c(t, S(t))\ dt + \partial_x c(t, S(t))[\alpha S(t)\ dt + \sigma S(t)\ dB(t)] + \tfrac{1}{2}\partial_x^2 c(t, S(t))\sigma^2 S^2(t)\ dt =$$
$$= [\partial_t c(t, S(t)) + \alpha S(t)\partial_x c(t, S(t)) + \tfrac{1}{2}\sigma^2 S^2(t)\partial_x^2 c(t, S(t))]\ dt +$$
$$+ \sigma S(t)\partial_x c(t, S(t))\ dB(t)$$

Next, let's compute the discounted option price's differential

$$d(e^{-rt}c(t, S(t))) = df(t, c(t, S(t))) =$$
$$= \partial_t f(t, c(t, S(t)))\ dt + \partial_x f(t, c(t, S(t)))\ dc(t, S(t)) +$$
$$+ \tfrac{1}{2}\partial_x^2 f(t, c(t, S(t)))\ dc(t, S(t))\ dc(t, S(t)) =$$
$$= -re^{-rt}c(t, S(t))\ dt + e^{-rt}\ dc(t, S(t)) =$$
$$= -re^{-rt}c(t, S(t))\ dt + e^{-rt}[(\partial_t c(t, S(t)) + \alpha S(t)\partial_x c(t, S(t)) + \tfrac{1}{2}\sigma^2 S^2(t)\partial_x^2 c(t, S(t)))\ dt +$$

$$+\sigma S(t)\partial_x c(t,S(t))\ dB(t)] =$$

$$= e^{-rt}[-rc(t,S(t)) + \partial_t c(t,S(t)) + \alpha S(t)\partial_x c(t,S(t)) +$$

$$+\tfrac{1}{2}\sigma^2 S^2(t)\partial_x^2 c(t,S(t))]\ dt + e^{-rt}\sigma S(t)\partial_x c(t,S(t))\ dB(t)$$

5.3 Equating the evolutions

In order to hedge a short position, we start with the initial capital $X(0)$, and invest in the stock and money markets in a way that for all $t \in [0,T]$ our portfolio value $X(t)$ agrees with the option value $c(t,S(t))$. One way to express this with the discounted prices is

$$e^{-rt}X(t) = e^{-rt}c(t,S(t))$$

or in differential form

$$d(e^{-rt}X(t)) = d(e^{-rt}c(t,S(t)))$$

Integrating both sides gives us

$$e^{-rt}X(t) - X(0) = e^{-rt}c(t,S(t)) - c(0,S(0))$$

so if we have $X(0) = c(0,S(0))$, the previous equation gives us the equality we strive for. But we already computed the differentials for the discounted values, so we just need to compare them. This leads us to the equation

$$\Delta(t)(\alpha - r)S(t)\ dt + \sigma\Delta(t)S(t)\ dB(t) \overset{!}{=}$$

$$\overset{!}{=} -rc(t,S(t)) + \partial_t c(t,S(t)) + \alpha S(t)\partial_x c(t,S(t)) + \tfrac{1}{2}\sigma^2 S^2(t)\partial_x^2 c(t,S(t))\ dt +$$

$$\sigma S(t)\partial_x c(t,S(t))\ dB(t)$$

By equating the $dB(t)$ terms we have

$$\Delta(t) = \partial_x c(t,S(t)) \ \ \forall\ t \in [0.T]$$

This is the *delta-hedging rule* (see Theorem 2.2.1), and we call $\partial_x c(t,S(t))$ the delta of the option. [1]
By equating the dt terms we get

$$(\alpha - r)S(t)\partial_x c(t,S(t)) =$$

$$= -rc(t,S(t)) + \partial_t c(t,S(t)) + \alpha S(t)\partial_x c(t,S(t)) + \tfrac{1}{2}\sigma^2 S^2(t)\partial_x^2 c(t,S(t)) \ \ \forall\ t \in [0,T]$$

Dropping out the terms $\alpha S(t)\partial_x c(t, S(t))$ from both sides we're left with

$$rc(t, S(t)) = \partial_t c(t, S(t)) + rS(t)\partial_x c(t, S(t)) + \tfrac{1}{2}\sigma^2 S^2(t)\partial_x^2 c(t, S(t))$$

So to summarize, we are searching for a continuous function $c(t, x)$ that is a solution to the *Black-Scholes-Merton partial differential equation*

$$\partial_t c(t, x) + rx\partial_x c(t, x) + \tfrac{1}{2}\sigma^2 x^2 \partial_x^2 c(t, x) = rc(t, x) \quad \forall t \in [0, T],\ x \geq 0$$

and satisfies the terminal condition

$$c(T, x) = \max(0, x - K)$$

If we find this function and start with the initial capital $X(0) = c(0, S(0))$ and use the hedge $\Delta_(t) = \partial_x c(t, S(t))$, then $X(t) = c(t, S(t)) \quad \forall t \in [0, T]$. By allowing $t \to T$ and considering the continuity of both $X(t)$ and $c(t, S(t))$, we conclude $X(T) = c(T, S(T)) = \max(0, S(T) - K)$, meaning that we successfully hedged a short position in the option.

Chapter 6

Other directions

Due to the extensive literature covering option pricing, we felt the need to include a chapter giving directions to other approaches, as well as different kinds of options (in this thesis we only considered European call options, a so-called *vanilla* option, but option trading is way broader: there are american, asian, barrier, lookback and other kinds of options).

6.1 Barrier options

Definition 6.1.1 *[8] A barrier option is a financial derivative contract that is activated (knocked in) or extinguished (knocked out) if the price of the underlying asset crosses a certain barrier.*

For example, an up-and-out call gives the holder the payout of a European call if the underlying doesn't breach the barrier before expiration. Another kind is a *double-barrier* option, where there may be a lower and upper bound for the underlying price, or a *two-dimensional barrier*, where the payoff is determined by one asset, while the barrier is determined by another.
The issue with pricing such an option is barrier monitoring: it's either done continuously or discretely. Continuous modeling allows analytical solutions (see [9], where an algorithm is developed for barrier option pricing in a one-dimensional Markov model by constructing an approximating continuous-time Markov chain).
In practice, most barriers are monitored discretely; in this case, pricing is not that easy due to three factors: (1) there are basically no closed formulas, (2) Monte Carlo simulations

take hours, or even days to produce a result, and (3) though the Central Limit Theorem states that the difference between continuously and discretely monitored barrier options should be small, but numerical computations have shown that even for large numbers the difference is significant.
For a numerical approach on discretely monitored barrier option pricing see [10].
To deal with the problem stated in (3), Brodie and Glasserman showed that discretely monitored barrier options can be approximated by continuous considerations. [11]

6.2 Exotic options

We also want to include a collection of different exotic options; according to [12], there are eleven categories when it comes to dividing exotic options into classes, namely
(1) packages - are basically equivalent to portfolios containing only European calls, possibly cash, and the underlying asset
(2) compound options - are options of the kind where the underlying itself is an option
(3) forward-start options - are a kind where the payment is made in the present, but the contracts themselves are only received in the future
(4) chooser options - are paid for in the present, but are later decided by the holder whether they are call or put options
(5) barrier options - see previous section
(6) lookback options - whose payoff not only depends on the price of the underlying at the maturity date, but also on the minimum or maximum price of the underlying asset during the life of the option
(7) asian options - also called average-price, because the payout depends on the average price of the underlying asset during the life of the option
(8) exchange options - exchange one asset for another
(9) currency translated options - whose underlying asset or strike price is denominated in a foreign currency at a prefixed or random exchange rate
(10) rainbow options - options on risky assets that cannot be interpreted as if they were a collection of options on one risky underlying asset
(11) binary options – options with binary and discontinuous payoff patterns

Pricing these often goes beyond the limitations of the Black-Scholes-Merton formula, and requires more sophisticated approaches. One of these is Lévy models. The idea here is that instead of the geometric Brownian motion, one considers a *Lévy process*. The reason for doing so is that Lévy processes capture empirical characteristics of real-life markets, such as jumps due to market shocks. We direct the reader to [13] for exotic option pricing under Lévy models, and to [14] for the precise discussion of Lévy processes.

Bibliography

[1] Shreve, Steven E. Stochastic calculus for finance II: Continuous-time models. Vol. 11. New York: springer, 2004.

[2] Durrett, Rick. "Probability: Theory and Examples." (2019).

[3] Doob, Joseph L. "What is a Martingale?." The American Mathematical Monthly 78.5 (1971): 451-463.

[4] Klenke, Achim. Probability theory: a comprehensive course. Springer Science Business Media, 2013.

[5] Laczkovich, Miklós, and Vera T. Sós. Real analysis I. Typotex, 2012

[6] Hurley, Nicholas S. "No-Arbitrage Option Pricing and the Binomial Asset Pricing Model." (2015).

[7] Shreve, Steven. Stochastic calculus for finance I: the binomial asset pricing model. Springer Science Business Media, 2005.

[8] Kou, Steven G. "On pricing of discrete barrier options." Statistica Sinica (2003): 955-964.

[9] Mijatović, Aleksandar, and Martijn Pistorius. "Continuously monitored barrier options under Markov processes." Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics 23.1 (2013): 1-38.

[10] Hong, Yicheng, Sungchul Lee, and Tianguo Li. "Numerical method of pricing discretely monitored Barrier option." Journal of Computational and Applied Mathematics 278 (2015): 149-161.

[11] Broadie, Mark, Paul Glasserman, and Steven Kou. "A continuity correction for discrete barrier options." Mathematical Finance 7.4 (1997): 325-349.

[12] Martinkute-Kauliene, Raimonda. "Exotic options: a chooser option and its pricing." Business, Management and Education 10.2 (2012): 289-301.

[13] Agliardi, Rossella. "A comprehensive mathematical approach to exotic option pricing." Mathematical Methods in the Applied Sciences 35.11 (2012): 1256-1268.

[14] Schilling, René L. "An introduction to Lévy and Feller processes." Lévy-type processes to parabolic SPDEs. Birkhäuser, Cham (2016).

Nyilatkozat

Alulírott Elter Ádám Gusztáv nyilatkozom, hogy szakdolgozatom elkészítése során az alább felsorolt feladatok elvégzésére a megadott MI alapú eszközöket alkalmaztam:

Writefull - teljes szöveg nyelvhelyesség ellenőrzésére
ChatGPT - python kód generálás a 15. és 17. oldalon található ábrák elkészítéséhez.

A felsoroltakon túl más MI alapú eszközt nem használtam.