On p-mean options in general, and Asian options in particular

Erik Broman

August 24, 2001

Abstract

This thesis concerns the evaluation process of p-mean options. This class of options includes the subclasses Asian options and Lookback options. The first of these subclasses is a very important tool for economical insurance, while the other presents a chance to gain large profits but is relatively expensive. Both of these are well known throughout the literature of financial mathematics, and many papers have been written on the subject of evaluating these. However, to the best of our knowledge none has used the Finite Element Method.

We present some results showing that the Finite Element Method can perform very well in some circumstances. It performs with high accuracy, even when using fewer degrees of freedom than comparable numerical methods ([19]).

We also present an approach to a similarity reduction of the fixed strike Asian call option. This approach is much less complex than the previously known approach due to Rogers and Shi [13], but has never been presented before as far as we know.

Acknowledgements

I would like to thank my supervisors Christer Borell and Mats G. Larson for their support and advice throughout the process of writing this thesis. I would also like to thank Endre Süli at the Oxford University Computing Laboratory for accepting me as a visiting student in Oxford and for the support I received in writing this thesis, especially the part concerning Streamline Diffusion.

Last but not least, I would like to thank my sister Malin for all the inspiration I have received throughout the years, and for the ticket to Australia.

Contents

1 Introduction

An option gives the holder the right, but not the obligation, to buy or sell a certain commodity for a price determined by an initially agreed formula at a settlement date or within a designated period. There are two main purposes of options. As a tool of speculation, and as a tool of insurance (or indeed as have been pointed out to us, as a tool for having fun). Take for example a share X with the price of \$1 at the date $t = 0$. Assume that an investor were to believe that the share is undervalued, and therefore that the price of the share is likely to rise within a certain period (say until $t = T$). Then the investor could buy a number of shares, and sell them later to earn money on his beliefs. If the share price at $t = T$ had risen to \$1.10, the investor would have made a profit of (disregarding the effects of interest rates, and transaction costs)

$$\frac{\text{the share price at t=T - the share price at t=0}}{\text{initial investment}} = \frac{1.10 - 1}{1} = 10\%.$$

If instead the investor would have bought a number of options with the payoff equal to the maximum of 0 and *the share value at t=T - \$1* for the price of \$ 0.05 (a reasonable price), the profit would have been

$$\frac{\text{the share price at t=T - \$ 1- initial investment}}{\text{initial investment}} = \frac{1.10 - 1 - 0.05}{0.05} = 100\%.$$

Note that, if the share value dropped to under 1\$ the investor would loose everything. The option described above is very suitable for gaining large profits compared to the option price.

As an example of an option suitable for insurance purposes, take for instance the Asian option, which in a certain variant pays the holder of the option the maximum of 0, and a strike price K minus the average of the underlying contingent A at expire. This payoff is commonly written as $\max(0, K - A)$. A company Y may have to make large investments in a foreign currency at a few time instances while continually selling a product in the same currency. As an example take a manufacturing company based in Sweden. They buy industrial robots from the USA, once every other year, but they sell their products back to the USA at a steady rate. This company is very vulnerable to fluctuations in the US dollar. To insure the company against changes in the value of the US dollar, the company could buy a number of the options mentioned above at the time of buying the robots. If the value of the dollar dropped during the following period, the value of the option would increase, making up for the decrease of profit from selling the manufactured product. Of course the options are not for free, so the company would have to decide what risk they would be willing to take, bearing in mind the cost of insurance.

The history of options is long and can be divided into two main parts, pre- and post- Black-Scholes. The first known example comes from Thales (624-547 BC), who lived in Miletus, Greece. Aristotele wrote:

> There is an anecdote of Thales the Milesian and his financial device, which involves a principle of universal application, but which is attributable to him on account of his reputation for wisdom. He was reproached for his poverty, which was supposed to show that philosophy was of no use. According to the story, he knew by his skill in the stars while it was yet winter that there would be a great harvest of olives in the coming year, so, having little money, he gave deposits for the use of all the olive presses in Chios and Miletus, which he hired at a low price because no one bid against him. When the harvest time came, and many wanted them all at once and of a sudden, he let them out at any rate which he pleased, and made a quantity of money. Thus he showed the world that philosophers can easily be rich if they like...

It is worth mentioning at this stage that astrology is not an accepted method in the field of option evaluation theory. Whether it is or not in other parts of financial theory, we can only speculate about. Another notable and amusing anecdote comes from 17th century Holland. About 1600, huge amounts of money were paid for tulips, and for tulip bulbs. In order to insure themselves against sudden drop in prices, the growers bought put options (i.e. options with the payoff max(0, *fixed strike-price – the tulip price*)), and the retailers bought call options (payoff max(0, *tulip price – fixed strike price*)) to insure themselves against steeply rising prices. However, the market eventually crashed in February 1637 following months of speculation resulting in outrageous prices, for tulip options.

Option trading has been a reality for a long time, even on official boards. But it was first after Fischer Black and Myron Scholes presented their consistent treatment of the subject [3], and a reliable valuation model for a number of simple options, that full-scale trading was allowed. It started in Chicago, on the Chicago Board Option Exchange in 1973, and has since spread to all financial corners of the world.

We would like to present some common terminology concerning options. A *put* option gives the holder the right to *sell* an asset, while a *call* option gives the holder the right to *buy* an asset. An Asian option in general depends not only on the underlying asset price, but also of the time average of this price. A typical example gives the holder the right to buy an asset for its average price over some period. Consistently, there exist Asian calls and Asian puts. When we talk about a simple call option, we are referring to the type of option with a payoff max(0, *the share value at expire date –*

fixed strike price), and similar for the put. These two kinds are commonly referred to as vanilla options.

In this thesis, we will study p-mean options and especially the form of Asian option mentioned in the example of the manufacturing company above. An Asian call option (or a fixed strike Asian option) is an option with the payoff $max(0, A - K)$, A for continuous average of the stock price and K for the strike price determined beforehand. The meaning of a p-mean option will become clear in the following sections, and Section 11 is entirely devoted to numerical results of p-mean options.

2 Fundamental Concepts

The model for the evolution of the stock value $S(t)$ used in this paper is standard in financial literature, and is commonly referred to as the Bachelier-Samuelson model. Bachelier was the first to have the idea that the stock price could be accurately described as a diffusion process, an idea which he presented in the article "Théorie de la spéculation"[1] [1], about year 1900. This point of view became popular first after Samuelson's "rediscovery" in 1965, where he presented an article in which the stock price evolution was modeled as a geometric Brownian motion with drift [15]. In this model the value of a bond is assumed to develop as

$$dB(t) = B(t)rdt$$

while the stock value is assumed to develop as

$$dS(t) = S(t)(\mu dt + \sigma dW(t)). \tag{1}$$

Here, $\alpha = \mu - \frac{\sigma^2}{2}$ is the assumed drift of the stock log-price, σ the volatility and $W(t)$ a normalized Wiener process. Let $\Omega_T = C[0,T]$, P be the corresponding Wiener measure and $W(t,\omega) = \omega(t), \quad \omega \in \Omega_T$. Then, $W(t)$ is a normalized Wiener process with respect to the probability space $(\Omega_T, \mathcal{F}_T, P)$, where $\mathcal{F}_T = \sigma(W(t); t \leq T)$. Define

$$\tilde{W}(t) = \frac{\mu - r}{\sigma} t + W(t). \tag{2}$$

Furthermore, by letting

$$P^a[A] = P[A - \{a\}] \text{ for } A \in \mathcal{F}_T$$

and

$$W^a(t) = W(t) - a(t)$$

it follows (see [4]) that $W^a(t)$ is a normalized Wiener process with respect to the probability space $(\Omega_T, \mathcal{F}_T, P^a)$. Let us now state a very important result in these financial environments.

Theorem 2.1 (Cameron-Martin) *Let $h \in L^2([0,T], \mathcal{B}([0,T]))$, and define $a \in \Omega_T$ as*

$$a(t) = \int_0^t h(\lambda) d\lambda \quad \forall t \in [0,T].$$

Then

$$dP^a(\omega) = e^{[\int_0^T h(t)dW(t)](\omega) - \frac{1}{2}\int_0^T h(t)^2 dt} dP(\omega). \tag{3}$$

[1] The greatness of his worked was not understood at the time he wrote the article, and his contribution was not recognized until after his death in 1946.

For a proof, see for example [4]. So by letting $h = -\frac{\mu - r}{\sigma}$ and $a = \int_0^t h(\lambda)d\lambda$ in (2) we know that there is a probability measure $\tilde{P}$ such that $\tilde{W}(t)$ is a normalized Wiener process with respect to $(\Omega_T, \mathcal{F}_T, \tilde{P})$ and by equation (3), P and $\tilde{P}$ are equivalent[2]. The probability measure $\tilde{P}$, the so-called "risk-neutral" measure is commonly denoted by Q. The definition of the process $\tilde{W}(t)$ yields

$$dS(t) = S(t)(rdt + \sigma d\tilde{W}(t)). \tag{4}$$

We will need the following three definitions.

Definition 2.2 (Progressively measurable) *If the mapping*

$$(t, \omega) \to h(t, \omega), \ (t, \omega) \in [0, T_0] \times \Omega_T$$

is $(\mathcal{B}[0, T_0] \times \mathcal{F}_{T_0})$-measurable for every fixed $T_0 \in [0, T]$, then the stochastic process $h = (h(t))_{0 \le t \le T}$ is said to be progressively measurable.

Here ω denotes a realization of the underlying Wiener process.

Definition 2.3 *If the stochastic process $h(t, \omega) = (h(t))_{0 \le t \le T}$ is progressively measurable, and*

$$\int_0^T | h(t) |^p \, dt < \infty \ \ a.s. \quad [P]$$

for some $p \in [1, \infty[$, then $h \in L_W^p[0, T]$.

Definition 2.4 *If the stochastic process $h(t, \omega) = (h(t))_{0 \le t \le T}$ is progressively measurable, and*

$$E[| h(t) |^p] < \infty$$

for some $p \in [1, \infty[$, then $h \in M_W^p[0, T]$.

Here W denotes the underlying Wiener process.

Consider a portfolio consisting of $h_s(t)$ stocks and $h_b(t)$ bonds at time t, where $h_s(t)$ and $h_b(t)$ are $\mathcal{F}_t$-measurable and with the restriction that $h_s S \in M_{\tilde{W}}^2[0, T]$, and $h_b \in L_W^1[0, T] = L_{\tilde{W}}^1[0, T]$. The value of the portfolio at time t equals

$$v(t, S(t)) = h_s(t)S(t) + h_b(t)B(t). \tag{5}$$

The portfolio strategy [3] is said to be self-financing if $v(t)$ fulfils

$$\begin{aligned} dv(t, S(t)) &= h_s(t)dS(t) + h_b(t)dB(t) \\ &= h_s(t)dS(t) + h_b(t)rB(t)dt, \ \ \forall \, t \in [0, T]. \end{aligned} \tag{6}$$

We will need the definition of a non-arbitrage portfolio-strategy.

[2]Two positive measures μ and ν on a measure space $(X, \mathcal{A})$ are said to be equivalent if $\mu(A) = 0 \Leftrightarrow \nu(A) = 0 \ \ \forall A \in \mathcal{A}$.

[3]A sequence of holdings in the stock and the bond, i.e. $(h_s(t), h_b(t))_{0 \le t \le T}$.

Definition 2.5 *A self-financing strategy is said to be free of arbitrage if, the two conditions* $v(0) = 0$ *and* $v(T) \geq 0$ *implies that* $v(T) = 0$ *a.s., i.e.* $P[v(T) = 0] = 1$.

The existence of the risk-neutral measure guarantees that every self-financing strategy is a non-arbitrage strategy. To see this, we state the following theorem were we use the risk-neutral measure Q.

Theorem 2.6 *Every self-financing portfolio strategy in the stock and bond is free of arbitrage.*

Proof. Let $(v(t), h_s(t), h_b(t))_{0 \leq t \leq T}$ be as above with $v(0) = 0$ and $v(T) \geq 0$. Define

$$X(t) = \frac{v(t)}{B(t)}$$

from which it follows that

$$d(B(t)X(t)) = h_s(t)dS(t) + h_b(t)dB(t) = h_s(t)dS(t) + \frac{B(t)X(t) - h_s(t)S(t)}{B(t)} dB(t).$$

Define

$$g(t) = \sigma \frac{h_s(t)S(t)}{B(t)}$$

and thus

$$h_s(t)dS(t) = \frac{g(t)B(t)}{\sigma}(rdt + \sigma d\tilde{W}(t)) = \frac{g(t)}{\sigma} dB(t) + g(t)B(t)d\tilde{W}(t).$$

Hence

$$d(B(t)X(t)) = g(t)B(t)d\tilde{W}(t) + X(t)dB(t)$$

and therefore

$$dX(t) = g(t)d\tilde{W}(t).$$

Now since $g(t) \in M^2_{\tilde{W}}[0, T]$,

$$X(T) = X(0) + \int_0^T g(t)d\tilde{W}(t) \Rightarrow X(0) = E^Q[X(T)].$$

To see that $E^Q[\int_0^T g(t)d\tilde{W}(t)] = 0$, consider the step function $f \in M^2_{\tilde{W}}[0, T]$

$$\int_0^T (f(t))^2 dt = \sum_{k=0}^{n-1} f^2(t_k)(t_{k+1} - t_k);$$

hence

$$\infty > E^Q[(\int_0^T f(t)d\tilde{W})^2] = \sum_{k=0}^{n-1} E^Q[f^2(t_k)](t_{k+1} - t_k);$$

and therefore $f(t_k,\omega) \in L^2(Q),\ k=0,\ldots,n-1$. Since if in general $Y(\omega) \in L^2(Q)$

$$(\int_{\Omega_T} \mid Y \mid dQ)^2 \le \int_{\Omega_T} dQ \int_{\Omega_T} \mid Y \mid^2 dQ = \int_{\Omega_T} \mid Y \mid^2 dQ < \infty,$$

by the use of Cauchy-Schwarz inequality, $Y(\omega) \in L^1(Q)$ and hence $f(t_k,\omega) \in L^1(Q),\ k=0,\ldots,n-1$, and therefore

$$\begin{aligned} E^Q[\int_0^T f(t)d\tilde{W}] &= \sum_{k=0}^{n-1} E^Q[f(t_k)(\tilde{W}(t_{k+1}) - \tilde{W}(t_k))] \\ &= \sum_{k=0}^{n-1} E^Q[f(t_k)]E^Q[(\tilde{W}(t_{k+1}) - \tilde{W}(t_k))] = 0, \end{aligned}$$

since $E^Q[f(t_k)] < \infty$. So from the approximation theorem stated below, we conclude that

$$E^Q[\int_0^T g(t)d\tilde{W}] = 0.$$

If $X(0) = v(0)/B(0) = 0$, it follows that $X(T) = 0$ a.s. [Q], from which we conclude that $v(T) = B(T)X(T) = 0$ a.s. [Q] and since [Q] and [P] are equivalent $P[v(T) = 0] = 1$

QED

Theorem 2.7 (Approximation theorem for $M^2_{\tilde{W}}(0,T)$) *Suppose $h \in M^2_{\tilde{W}}[0,T]$. There exists step functions $h_n \in M^2_{\tilde{W}}, n \in N$, such that*

$$\lim_{n\to\infty} E^Q[\int_0^T (g_n(t) - g(t))^2 dt] = 0$$

For a proof see [11].

From the above it follows that we have no arbitrage in the Black-Scholes model. It is worth mentioning that it is easy to generalize Theorem 2.6 to several dimensions (i.e. several underlying stocks).

Another very important and often used tool in financial mathematics is Itô's lemma.

Lemma 2.8 (Itô) *Let $u(t,x_1,\ldots,x_m)$ be one time continuously differentiable in $t \in [0,T]$ and two times continuously differentiable in $x_1,\ldots,x_m \in$* **R**. *Suppose:*

$$dX_i(t) = a_i(t) + \sum_{k=1}^{n} b_{ik}(t)dW_k(t)$$

where $W_1(t),\ldots,W_n(t)$ are independent Wiener processes, and $a_i(t) \in L^1_W[0,T],\ b_{ik}(t) \in L^2_W[0,T]\ \ \forall i \in (1,\ldots,m), k \in (1,\ldots,n)$. By letting

$$X(t) = (X_1(t),\ldots,X_m(t))$$

it follows that

$$du(t,X(t)) = \frac{\partial u}{\partial t}(t,X(t))dt + \sum_{i=1}^{m} \frac{\partial u}{\partial x_i}(t,X(t))dX_i(t) + \frac{1}{2}\sum_{i,j=1}^{m} \frac{\partial^2 u}{\partial x_i \partial x_j}(t,X(t))dX_i(t)dX_j(t). \quad (7)$$

Furthermore, by definition

$$(dt)^2 = 0, dt \quad dW_i(t) = 0,$$
$$dW_i(t)dW_j(t) = \begin{cases} dt & i = j \\ 0 & i \neq j \end{cases}$$

Lemma 2.9 (Coefficient matching) *If* $a_k \in L_W[0,T],\ b_k \in L^2_W[0,T],\ k = 1,2$ *and*

$$a_1(t)dt + b_1(t)dW(t) = a_2(t)dt + b_2(t)dW(t)$$

then $a_1 = a_2$ *and* $b_1 = b_2$.

Proof. Set $a = a_1 - a_2,\ b = b_1 - b_2$, Then

$$X(t) - X(0) = \int_0^t a(\lambda)d\lambda + b(\lambda)dW(\lambda) = 0.$$

Hence $dX(t) = 0$, and $dX^2(t) = 0$. But $dX^2(t) = 2X(t)dX(t) + b^2(t)dt$ and we get $b^2(t)dt = 0$. From this we conclude that $a = b = 0$, and therefore $a_1 = a_2$ and $b_1 = b_2$.

QED

We will now derive the so called Black-Scholes pricing formula for an option with a payoff function of the type

$$f(S(T_1), \ldots, S(T_n)),$$

based on the following definition.

Definition 2.10 *Let* $\mathcal{P}$ *be the set of all continuous non-negative functions* f *for which there exists a* $C > 0$ *such that* $\sup\{e^{-C|x|}f(e^x); x \in \mathbf{R}\} < \infty$. *A simple European option with the payoff* $f(S(T)), f \in \mathcal{P}$, *has at time* $t < T$ *the theoretical value*

$$v(t,S(t)) = e^{-r(T-t)}E[f(se^{(r-\sigma^2/2)(T-t)+\sigma W(T-t)})]_{s=S(t)}.$$

Again we let

$$a = \frac{r-\mu}{\sigma} t, \quad 0 \le t \le T.$$

By the use of Cameron-Martins theorem we conclude that

$$dQ(\omega) = e^{\frac{r-\mu}{\sigma} W(T) - \frac{1}{2}\left(\frac{r-\mu}{\sigma}\right)^2 T} dP(\omega),$$

which implies that

$$dP(\omega) = e^{-\frac{r-\mu}{\sigma} W(T) + \frac{1}{2}\left(\frac{r-\mu}{\sigma}\right)^2 T} dQ(\omega).$$

Recall that

$$\tilde{W}(t) = W(t) - a(t)$$

is a normalized Wiener process with respect to the probability space $(\Omega_T, \mathcal{F}_T, Q)$.

Theorem 2.11 *Assume that $f \in C(\mathbf{R}^n)$ and that $t \le T_1 \le \cdots \le T_n \le T$. It follows that*

$$\begin{aligned} &E^Q[f(S(T_1), \ldots, S(T_n)) | \mathcal{F}_t] \\ &\quad = E[f(se^{(r-\frac{\sigma^2}{2})(T_1-t)+\sigma(W_{T_1}-W_t)}, \ldots, se^{(r-\frac{\sigma^2}{2})(T_n-t)+\sigma(W_{T_n}-W_t)})]_{s=S(t)}. \end{aligned}$$

Proof.

$$\begin{aligned} &E^Q[f(S(T_1), \ldots, S(T_n)) | \mathcal{F}_t] \\ &\quad = E^Q[f(S(t)e^{(\mu-\frac{\sigma^2}{2})(T_1-t)+\sigma(W_{T_1}-W_t)}, \ldots, S(t)e^{(\mu-\frac{\sigma^2}{2})(T_n-t)+\sigma(W_{T_n}-W_t)}) | \mathcal{F}_t] \\ &\quad = E^Q[f(S(t)e^{(r-\frac{\sigma^2}{2})(T_1-t)+\sigma(\tilde{W}_{T_1}-\tilde{W}_t)}, \ldots, S(t)e^{(r-\frac{\sigma^2}{2})(T_n-t)+\sigma(\tilde{W}_{T_n}-\tilde{W}_t)}) | \mathcal{F}_t] \\ &\quad = E^Q[f(se^{(r-\frac{\sigma^2}{2})(T_1-t)+\sigma(\tilde{W}_{T_1}-\tilde{W}_t)}, \ldots, se^{(r-\frac{\sigma^2}{2})(T_n-t)+\sigma(\tilde{W}_{T_n}-\tilde{W}_t)})] \\ &\quad = E[f(se^{(r-\frac{\sigma^2}{2})(T_1-t)+\sigma(W_{T_1}-W_t)}, \ldots, se^{(r-\frac{\sigma^2}{2})(T_n-t)+\sigma(W_{T_n}-W_t)})]. \end{aligned}$$

QED

By Definition 2.10

$$v(T_{n-1}) = e^{-r(T_n - T_{n-1})} E^Q[f(S(T_1), \ldots, S(T_n)) | \mathcal{F}_{T_{n-1}}],$$

which is a function of the type $g(S(T_1), \ldots, S(T_{n-1}))$, where g is non-negative and continuous. By considering g to be the payoff of an (fictitious) option with maturity T_{n-1}, we conclude that

$$v(T_{n-2}) = e^{-r(T_{n-1} - T_{n-2})} E^Q[g(S(T_1), \ldots, S(T_{n-1})) | \mathcal{F}_{T_{n-2}}].$$

Set $Y = f(S(T_1), \ldots, S(T_n))$. By induction, we conclude that

$$\begin{aligned} v(t) &= e^{-r(T_1-t)} E^Q[e^{-r(T_2-T_1)} E^Q[\cdots e^{-r(T_n-T_{n-1})} E^Q[Y | \mathcal{F}_{T_{n-1}}] \cdots | \mathcal{F}_{T_1}] | \mathcal{F}_t] \\ &= e^{-r(T_n-t)} E^Q[Y | \mathcal{F}_t], \end{aligned}$$

where we used the Tower Property of Conditional Expectation: If $\mathcal{H}$ is a sub-σ-algebra of $\mathcal{G}$, then $E[E[X|\mathcal{G}]\mathcal{H}] = E[X|\mathcal{H}]$. This result motivates the following definition

Definition 2.12 *The value $v(t)$ at time t of a European option with maturity date T and payoff function $Y \in L^2(\Omega_T, Q), Y \geq 0,$ equals*

$$v(t) = e^{-r(T-t)} E^Q[Y|\mathcal{F}_t], \quad \forall t \in [0, T].$$

This is commonly referred to as the Black-Scholes pricing formula.

The above formalism gives rise to the possibility of calculating the option value as an expectation. By simulating the trajectories of the Wiener processes involved in determining Y a large number of times, we can easily calculate an approximate value of this expectation. This method is commonly referred to as the Monte-Carlo method and we will use this method later on.

3 Average options

Armed with the theory from the previous section, we will now derive valuation equations for a certain class of options. A subset of these will be the so called p-mean options. To begin with set

$$Z(t) = \int_0^t g(S(\lambda), \lambda) d\lambda$$

so that

$$dZ(t) = g(S(t), t)dt.$$

Consider an option with the payoff function

$$f(S(T), Z(T)).$$

To evaluate an option of this form, we assume that the value can be written in the form[4]

$$v(t) = v(t, S(t), Z(t)). \tag{8}$$

We recall that by (4),

$$dS(t) = S(t)(rdt + \sigma d\tilde{W}(t)).$$

Put

$$X(t) = (S(t), Z(t)).$$

We will sometimes write

$$S(t) = s, \quad Z(t) = z,$$

and consequently

$$x = (s, z).$$

Now, applying Itô's lemma to equation (8), results in

$$\begin{aligned} dv(t, X(t)) &= v_t'(t, X(t))dt + v_s'(t, X(t))dS(t) + v_z'(t, X(t))dZ(t) \qquad (9) \\ &\quad + \frac{1}{2}\left(v_{ss}''(t, X(t))(dS(t))^2 + 2v_{sz}''(t, X(t))dS(t)dZ(t)\right. \\ &\quad \left. + v_{zz}''(t, X(t))(dZ(t))^2\right) \\ &= v_t'(t, X(t))dt + v_s'(t, X(t))dS(t) + v_z'(t, X(t))g(S(t), t)dt \\ &\quad + \frac{S^2(t)\sigma^2}{2} v_{ss}''(t, X(t))dt. \end{aligned}$$

So let $v(t, X(t))_{0 \le t \le T}$ be the value process of a self-financing strategy, and by equations (6) and (9), coefficient matching yields

$$h_s(t) = v_s'(t, X(t))$$

[4] See for example [18].

and

$$h_b(t)rB(t) = v'_t(t, X(t)) + v'_z(t, X(t))g(S(t), t) + \frac{S^2(t)\sigma^2}{2} v''_{ss}(t, X(t)).$$

From equation (5) it follows that

$$\begin{aligned} v(t, X(t)) &= S(t)v'_s(t, X(t)) + \frac{1}{r}(v'_t(t, X(t)) \\ &\quad + v'_z(t, X(t))g(S(t), t) + \frac{S^2(t)\sigma^2}{2} v''_{ss}(t, X(t))), \end{aligned}$$

from which we conclude that (omitting arguments of v)

$$v'_t + rsv'_s + g(s,t)v'_z + \frac{s^2\sigma^2}{2} v''_{ss} - rv = 0 \ , s > 0, t < T. \tag{10}$$

This equation will be used together with the final condition

$$v(T, s, z) = f(T, s, z),$$

appropriate to the option of interest. In the following sections, we will use the PDE formulation of evaluating option prices. We will then solve these PDE's using the Finite Element Method and compare the results with the option prices obtained by others and by using the Monte Carlo method.

4 Dimensionality reduction

Throughout the literature, different suggestions and approaches (see for example [13]) have been made to reduce the dimensionality of equation (10). We have chosen the simple method of change of variables, to find possible similarity reductions. Here our work have been concerning p-mean options, with a payoff function equal to either

$$f(T,S(T),Z(T)) = \max(Z(T)^{1/p} - K, 0) \quad (\text{“fixed strike”})$$

or

$$f(T,S(T),Z(T)) = \max(Z(T)^{1/p} - S(T), 0) \quad (\text{“floating strike”})$$

where

$$\begin{cases} Z(T) &= \int_0^T g(S(\lambda),\lambda)d\lambda \\ g(S(t),t) &= S^p(t)\rho(t). \end{cases}$$

An approach with

$$v(t,s,z) = su(t,y), \quad y = \frac{z}{s}$$

is known to reduce the dimensionality of the floating strike Asian call option, i.e., with the payoff

$$f(T,S(T),Z(T)) = \max(Z(T) - S(T), 0),$$

(see [18]). By using the transformation $v(t,s,z) = h(s)u(t,y), \quad y = y(s,z)$ we can write equation (10) as

$$u_t' + u(rs\frac{h_s'}{h} + \frac{\sigma^2 s^2}{2}\frac{h_{ss}''}{h} - r) + u_y'(rsy_s' + gy_z' + \sigma^2 s^2 \frac{h_s'}{h} y_s' + \frac{\sigma^2 s^2}{2} y_{ss}'') + \frac{\sigma^2 s^2}{2}(y_s')^2 u_{yy}'' = 0. \quad (11)$$

Consider for instance the payoff

$$f(T,S(T),Z(T)) = \max(Z(T) - K, 0), \quad Z(t) = \int_0^t S(\lambda)\rho(\lambda)d\lambda,$$

so here $g(S(t),t) = S(t)\rho(t)$. We make the ansatz

$$v(t,s,z) = su(t,y) = h(s)u(t,y), \quad y = \frac{z-K}{s}.$$

Plugging this into equation (11) gives us

$$\begin{aligned} & u_t' + u(rs\frac{1}{s} + \frac{\sigma^2 s^2}{2}\frac{0}{s} - r) \\ & \quad + u_y'(rs\frac{-y}{s} + s\rho(t)\frac{1}{s} + \sigma^2 s^2 \frac{1}{s}\frac{-y}{s} + \frac{\sigma^2 s^2}{2}\frac{2y}{s^2}) \\ & \quad + \frac{\sigma^2 s^2}{2}(\frac{-y}{s})^2 u_{yy}'' \\ & \quad = u_t' + u_y'(\rho(t) - ry) + \frac{y^2\sigma^2}{2}u_{yy}'' = 0, \end{aligned} \quad (12)$$

with the corresponding final condition

$$v(T,s,z) = s\max(\frac{z-K}{s},0) = s\max(y,0) \Rightarrow u(T,y) = \max(y,0).$$

For an alternative, although a bit more tedious derivation of this formula, see [13].

To sum up this specific example, we notice that for a discretely sampled Asian option, $\rho(t)$ will contain point masses. So by allowing ourselves for a moment to regard $\rho(t)$ as being made up by Dirac "functions", and setting[5]

$$\begin{aligned} q(t) &= \int_t^T \rho(\lambda)d\lambda \\ w &= q(t) + y \\ u(t,y) &= \mu(t,w) \end{aligned}$$

equation (12) becomes

$$\begin{aligned} & u_t' + u_y'(\rho(t) - ry) + \frac{y^2\sigma^2}{2}u_{yy}'' \\ &= \mu_t' - \rho(t)\mu_w' + \mu_w'(\rho(t) - r(w - q(t))) + \frac{((w-q(t))\sigma)^2}{2}\mu_{ww}'' \\ &= \mu_t' + r(q(t) - w)\mu_w' + \frac{((q(t)-w)\sigma)^2}{2}\mu_{ww}'' = 0. \end{aligned} \quad (13)$$

Written in this form, we can solve a PDE with numerical PDE-methods to determine the value of discretely sampled Asian options. Let $\nu(t)$ be a function without Dirac-pulses in the interval $]\Theta, T]$, $0 < \Theta < T$, such that for $a \in]\Theta, T]$, $\rho = \nu + \delta_a$. The price of the option must be continuous with regard to a, which follows from Definition 2.12 and the Dominated Convergence theorem. Hence the final condition $\mu(T,w) = \max(w,0)$ is valid *even if* ρ *has a Dirac-pulse in* $t = T$*!*

Consider now the payoff

$$f(S(T), \int_0^T g(S(\lambda),\lambda)d\lambda) = \max(0, (\frac{1}{T}\int_0^T g(S(\lambda),\lambda)d\lambda)^{1/p} - S(T))$$

with

$$g(s,t) = s^p,$$

and

$$Z(t) = \int_0^T g(S(\lambda),\lambda)d\lambda.$$

Make the ansatz

$$v(t,s,z) = su(t,y), \quad y = \frac{z}{s^p}.$$

[5]As suggested by Večeř, [17].

Observe that

$$\begin{cases} y'_z = \frac{1}{s^p} \\ \frac{h'_s}{h} = \frac{1}{s} \end{cases}$$

and that

$$y'_s = -\frac{pz}{s^{p+1}} = -\frac{py}{s}.$$

Differentiate this to deduce that,

$$y''_{ss} = \frac{py}{s^2} - \frac{p}{s}y'_s = \frac{py}{s^2} + \frac{p^2}{s^2}y$$

and substitute these results into equation (11) to obtain

$$\begin{aligned}
& u'_t + u(rs\frac{h'_s}{h} + \frac{\sigma^2 s^2}{2}\frac{h''_{ss}}{h} - r) \\
& \quad +u'_y(rsy'_s + gy'_z + \sigma^2 s^2 \frac{h'_s}{h}y'_s + \frac{\sigma^2 s^2}{2}y''_{ss}) + \frac{\sigma^2 s^2}{2}(y'_s)^2 u''_{yy} \\
& = u'_t + u(rs\frac{1}{s} + \frac{\sigma^2 s^2}{2}\frac{0}{s} - r) \\
& \quad +u'_y(rs\frac{-py}{s} + s^p\frac{1}{s^p} + \sigma^2 s^2 \frac{1}{s}\frac{-py}{s} + \frac{\sigma^2 s^2}{2}(\frac{py}{s^2} + \frac{p^2}{s^2}y)) \\
& \quad +\frac{\sigma^2 s^2}{2}(-\frac{py}{s})^2 u''_{yy} \\
& = u'_t + u'_y(-pry + 1 - p\sigma^2 y + \frac{\sigma^2}{2}(py + p^2 y)) + \frac{\sigma^2}{2}(-py)^2 u''_{yy} \\
& = u'_t + u'_y(-pry + 1 + \frac{\sigma^2 y}{2}(p^2 - p)) + \frac{\sigma^2}{2}(py)^2 u''_{yy} = 0.
\end{aligned}$$

The payoff will be

$$su(T, y) = s\max(0, (\frac{z}{s^p})^{1/p} - 1) = s\max(0, y^{1/p} - 1)$$

and hence

$$u(T, y) = \max(0, y^{1/p} - 1).$$

So for the general p-mean option with floating strike, we only have to solve a PDE in one spatial dimension. Furthermore, by letting p tend to infinity it should be possible to find a PDE for the floating strike lookback option.

5 The FEM approach for solving the value governing equation

Consider equation (10) with $g(s,t) = s^p$

$$v_t' + rsv_s' + s^p v_z' + \frac{s^2\sigma^2}{2} v_{ss}'' - rv = 0 \ , s > 0, t < T, \tag{14}$$

and write this in the form

$$-v_t' - \nabla \cdot (a(x)\nabla v) + b(x) \cdot \nabla v + c(x)v = 0, \ \ x = (s,z), \tag{15}$$

where

$$a(x) = [\begin{array}{cc} \frac{\sigma^2 s^2}{2} & 0 \\ 0 & 0 \end{array}], \quad b(x) = -[(r-\sigma^2)s \ s^p] \ \text{ and } c(x) = r.$$

Multiplying equation (15) by a function $\varphi(t,x)$ and then integrating over time and over a suitable bounded area of s and z denoted by Ω, we obtain

$$\begin{aligned} &\int_0^T \int_\Omega (-v_t' - \nabla \cdot (a(x)\nabla v) + b(x) \cdot \nabla v + c(x)v)\varphi d\Omega dt \\ &\quad = \int_0^T \int_\Omega (-v_t'\varphi + (a(x)\nabla v) \cdot \nabla\varphi + (b(x) \cdot \nabla v)\varphi + c(x)v\varphi) d\Omega dt \\ &\quad\quad - \int_0^T \int_\Gamma \varphi(a(x)\nabla v) \cdot \vec{n} d\Gamma = 0, \end{aligned} \tag{16}$$

using the Gauss divergence theorem. The boundary of Ω is denoted here by Γ, and $\vec{n}$ is the from Ω outward oriented unit normal.

The so called weak solution to equation (14), is to find $v \in W$, where

$$W = L^2([0,T]) \times H_1(\Omega), \ \ H_1(\Omega) = \{v : \int_\Omega (v^2 + |\nabla v|^2) d\Omega < \infty)\}$$

such that

$$\begin{aligned} &\int_0^T \int_\Omega (-v_t'\varphi + (a(x)\nabla v) \cdot \nabla\varphi + (b(x) \cdot \nabla v)\varphi + c(x)v\varphi) d\Omega dt \\ &\qquad\qquad - \int_0^T \int_\Gamma \varphi(a(x)\nabla v) \cdot \vec{n} d\Gamma = 0 \ \ \forall \varphi \in W. \end{aligned}$$

However, to find the weak solution is not an altogether easy task, and we will therefore seek only an approximate solution. Partition the time interval $[0,T]$ into $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ where $I_n = [t_{n-1}, t_n], k_n = t_n - t_{n-1}$. For each I_n, let $\psi_{n-1}(t) = \frac{t_n - t}{k_n}$ and $\psi_n(t) = \frac{t - t_{n-1}}{k_n}$ for $t \in I_n$.

Subdivide the relevant area Ω into triangles of suitable size, and denote this triangulation $\mathcal{T}$. A triangle of this triangulation will be denoted κ, hence

$\mathcal{T} = \{\kappa\}$. Furthermore, let $\mathcal{P}_k \subset H_1(\Omega)$ denote the space of all continuous piecewise polynomials of degree k defined on $\mathcal{T}$[6]. Now, define

$$W_n^1 = \{w(t,x) : w(t,x) = \psi_{n-1}(t)v_{n-1}(x) + \psi_n(t)v_n(x), \\ v_{n-1}, v_n \in \mathcal{P}_k, \ \ (t,x) \in I_n \times \Omega\}$$

and

$$W_n^0 = \{w(t,x) : w(t,x) = t^0\phi(x), \ \ \phi \in \mathcal{P}_k, \ \ (t,x) \in I_n \times \Omega\}$$

and finally

$$W^i = \{\varphi : \varphi|_{I_n\times\Omega} \in W_n^i\}, \ \ i = 0, 1.$$

A possible finite element approximation to the formulation of equation (14), would be to find $v \in W^1$ such that

$$\int_0^T \int_\Omega (-v_t'\varphi + (a(x)\nabla v)\cdot\nabla\varphi + (b(x)\cdot\nabla v)\varphi + c(x)v\varphi)d\Omega dt \\ - \int_0^T \int_\Gamma \varphi(a(x)\nabla v)\cdot\vec{n}d\Gamma = 0, \ \ \forall\varphi \in W^0. \tag{17}$$

There are several possible ways to make other finite element approximations to equation (14). One possibility is to use higher order approximation in time, and define the spaces W^q analogously to W^0 and W^1. Another is to use quadrilateral elements instead of triangular in the subdivision of Ω.

For the type of equation that we are trying to solve, the domain of influence is localized in such a manner that the boundary do not play much role for the solution within the computational domain. What boundary conditions are enforced is therefore of not much significance to us, because we are looking for the solution at a point within the domain[7]. The finite element method requires us to impose a boundary condition, and by the argument above, we can choose almost whatever we like. We therefore choose a homogeneous Neumann condition on Γ, rendering the last term of equation (17) conveniently zero. This localized dependency is fortunate, because there are no financially motivated boundary conditions in this context (see Zvan et al. [19]).

We have that $\mathcal{P}_k = \text{span}\{\phi_1, \phi_2, \ldots, \phi_{M(\mathcal{T})}\}$, where $\phi_i, i = 1, \ldots, M(\mathcal{T})$, are orthonormal, have small support, and are commonly referred to as the nodal basis for $\mathcal{P}_k$[8]. Let us now expand $v_n(x)$ in this nodal basis

$$v_n(x) = \sum_{i=0}^{M(\mathcal{T})} V_{n,i}\phi_i(x),$$

[6] k will be equal to 1 or 2 in the numerics later on.
[7] For a further discussion about this see [19].
[8] See [7].

and also use the ϕ_i as test functions, observing that these are indeed piecewise constant in time since $\varphi_i = t^0\phi_i = \phi_i$. Inserting this into equation (17) yields (with $V_n = (V_{n,0}, V_{n,1}, \ldots, V_{n,M(\mathcal{T})})$)

$$-(V_n - V_{n-1})M + \frac{k_n}{2}(V_n + V_{n-1})S_1 + \frac{k_n}{2}(V_n + V_{n-1})S_2 + r(V_n + V_{n-1})M\frac{k_n}{2} = 0$$

where

$$M = (\phi_j, \phi_i), \quad S_1 = (b(s,z)\cdot\nabla\phi_j, \phi_i), \quad S_2 = (a(s,z)\nabla\phi_j, \nabla\phi_i).$$

After rearranging, this becomes

$$V_{n-1}(M(1 + r\frac{k_n}{2}) + \frac{k_n}{2}S_1 + \frac{k_n}{2}S_2) = V_n(M(1 - r\frac{k_n}{2}) - \frac{k_n}{2}S_1 - \frac{k_n}{2}S_2),$$

We look upon this as a system of linear equations that we can easily solve for each time step. Our final condition will determine the last $V_n = V_N$ for us, and we can then successively determine the value of the option for $t = t_n$, until we reach present date, $t = 0$.

6 Duality

For these calculations, only a simple error analysis is done. From the 3D-solution on Ω, we are really only interested in one point. Hence, it is important to know which is the area that influences the solution at this point, and spend our computational efforts here rather than somewhere else. By considering the adjoint operator $\mathcal{L}^*$, defined by $\langle \mathcal{L}v, \Psi\rangle = \langle v, \mathcal{L}^*\Psi\rangle$, ($\langle\cdot,\cdot\rangle$ denoting $\int_0^T\int_\Omega$) we can formulate the dual problem to equation (14)[9]. Remembering that

$$\mathcal{L} = \frac{\partial}{\partial t} + rs\frac{\partial}{\partial s} + s^p\frac{\partial}{\partial z} + \frac{s^2\sigma^2}{2}\frac{\partial^2}{\partial s^2} - r,$$

it is easy to conclude simply by doing the calculations leading to $\langle v, \mathcal{L}^*\Psi\rangle$ from $\langle \mathcal{L}v, \Psi\rangle$, that the dual problem must be

$$\begin{aligned} \mathcal{L}^*\Psi = -\Psi'_t - (r-\sigma^2)s\Psi'_s - s^p\Psi'_z & & \\ +\frac{\partial}{\partial s}(\frac{\sigma^2 s^2}{2}\Psi'_s) - (2r-\sigma^2)\Psi &=& 0 \\ \Psi(t=0,s,z) &=& \delta_{s_\alpha,z_\alpha} \\ \Psi(t,s,z) &=& 0, \quad t>0, \quad (s,z)\in\Gamma. \end{aligned} \tag{18}$$

Now, to see why this is relevant, we multiply the equation above by a function e, and transfer the derivatives of Ψ, onto e

$$\begin{aligned} &\int_0^T\int_\Omega(-\Psi'_t - (r-\sigma^2)s\Psi'_s - s^p\Psi'_z \\ &\quad +\frac{\partial}{\partial s}(\frac{\sigma^2s^2}{2}\Psi'_s) - (2r-\sigma^2)\Psi)e\,dtd\Omega \\ &= \int_\Omega(\Psi(0)e(0) - \Psi(T)e(T))d\Omega \\ &\quad + \int_0^T\int_\Omega(\Psi e'_t + (r-\sigma^2)\Psi\frac{\partial}{\partial s}(se) + \Psi\frac{\partial}{\partial z}s^pe \\ &\quad -(\frac{\sigma^2s^2}{2}\Psi'_s)\frac{\partial}{\partial s}(e) - (2r-\sigma^2)\Psi e)dtd\Omega \\ &= \int_\Omega(\Psi(0)e(0) - \Psi(T)e(T))d\Omega \\ &\quad + \int_0^T\int_\Omega\Psi(e'_t + rse'_s + s^pe'_z + \frac{s^2\sigma^2}{2}e''_{ss} - re)dtd\Omega \\ &= \int_\Omega(\Psi(0)e(0) - \Psi(T)e(T))d\Omega + \int_0^T\int_\Omega(\Psi\mathcal{L}e)dtd\Omega = 0. \end{aligned}$$

Choose $e = v - V$, where v solves equation (14) combined with the relevant final condition for the option of interest, and V is the approximate

[9] For a more extensive discussion of duality see [7].

solution, i.e. the solution to equation (17) (e for error). Since $v(T)$ is known, $e(T) = 0$, and this in combination with $\Psi(t = 0, s, z) = \delta_{s_\alpha, z_\alpha}$, makes us conclude that

$$
\begin{aligned}
e(t = 0, s_\alpha, z_\alpha) &= -\int_0^T \int_\Omega (\Psi \mathcal{L} e) dt d\Omega \\
&= -\int_0^T \int_\Omega (\Psi \mathcal{L}(v - V)) dt d\Omega = \int_0^T \int_\Omega (\Psi \mathcal{L} V) dt d\Omega. \qquad (19)
\end{aligned}
$$

We now have a closed expression for the error in a point (s_α, z_α) at time $t = 0$, expressed in terms of the solution to the dual problem (Ψ) and $\mathcal{L}V$. The magnitude of $\mathcal{L}V$ represents the error we make by solving equation (17) rather than equation (14).

7 Results and conclusions drawn from the dual problem

During the computations $\mathcal{P}_k$ was chosen sometimes as $\mathcal{P}_1$ and sometimes as $\mathcal{P}_2$, corresponding to cG1 and cG2 respectively (continuous Galerkin method of order 1 and 2). In general the latter case gave more accurate results and was therefore the preferred method. Some of the results in this and following sections were derived with the cG1-method as well as with the cG2-method, but we will only present the results for the cG2-method since this is the only one of the two used throughout all numerics.

Even though the FEM-software developed to solve equation (10) with $g(t, s) = s^p$, was designed to handle arbitrary p, most of the results in this thesis is for $p = 1$, simply because this is a well studied case, and there is a lot of literature to compare with. We will state some results for $p \neq 1$ as well.

The solution v of equation (17) is defined on a whole surface (Ω) rather than in a specific point (s_α, z_α). We can easily draw the solution surfaces using MATLAB's PDE-toolbox, which was also used to initiate and refine the mesh used for the calculations. The final condition $v(T) = \max(z - K, 0)$ (here $p = 1$), will generate two planes as the solution surface for $t = T$ (see the upper left frame of Figure 1). These two planes will thereafter propagate through the 3D-space as seen throughout the rest of the frames of Figure 1. It would seem from these figures that a mesh of about $20 - 100$ nodes would suffice. However this is not true, a careful study of the solution-curves for $t \neq T$ reveals that the two-plane structure somewhat dissolves and especially so for the points of interest (i.e. $(S(0) = s_0, Z(0) = z_0 = 0)$). On the contrary, the solution requires, in certain areas, a very fine computational mesh.

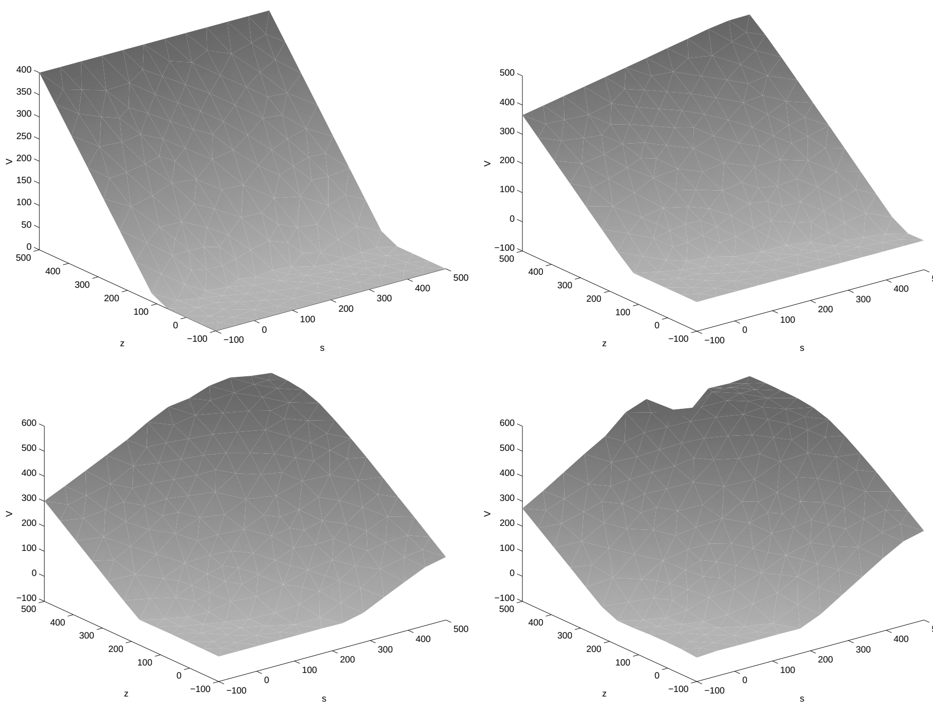

Figure 1: Above left shows $V(t = 0)$, above right $V(t = T/4)$, below left $V(t = 3T/4)$, and below right shows $V(t = T)$.

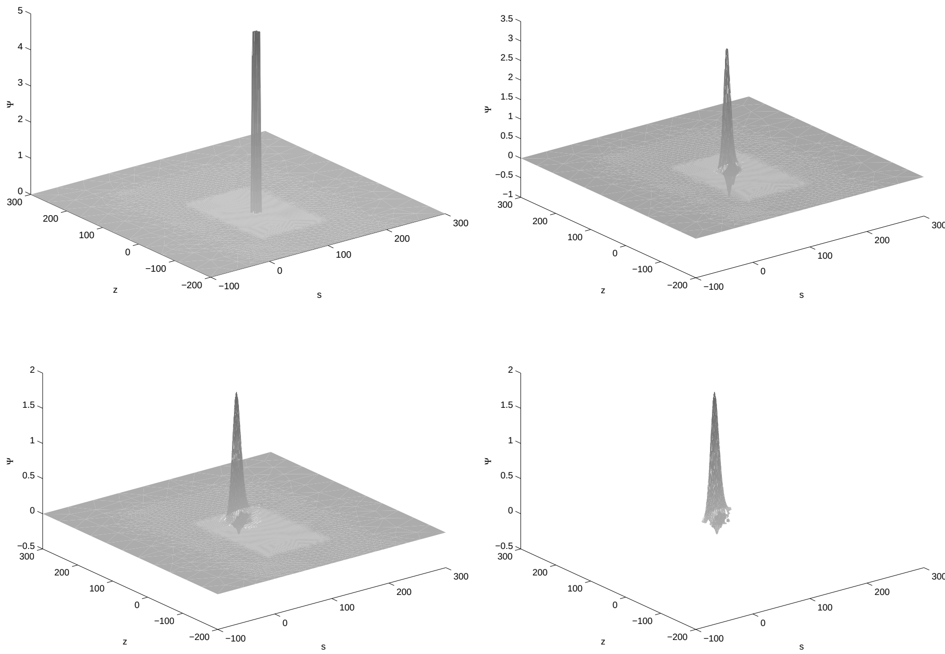

Figure 2: Above left, is showing $\Psi(t=0)$, above right $\Psi(t=T/2)$, below left shows $\Psi(t=T)$, and below right shows $\Psi(t=T)$ as well, but only values of Ψ such that $\Psi > \frac{1}{100}$

Consider equation (19). The error depends heavily on Ψ and $\mathcal{L}V$, although of course the choice of Ω and model errors are significant as well. We might not know much about $\mathcal{L}V$ without actually calculating it, but we do know that for an exact solution it would be equal to zero. However since we are only approximating the solution numerically, $\mathcal{L}V$ will in general not be equal to zero. But, from the equation it is easy to conclude that if Ψ is small or even zero for some parts of Ω, the error $e(t=0, s_\alpha, z_\alpha)$ does not depend on the magnitude of the approximation error $|\mathcal{L}(v-V)|$ in these same areas. For all practical purposes this means that wherever Ψ is small, we should not waste any computational effort to minimize $|\mathcal{L}(v-V)|$. And, of course, wherever Ψ is large, we should be extremely careful to make our solution as exact as possible. We shall therefore use a coarse mesh where Ψ is small and a fine where Ψ is large. In Figure 2 we can see the result of the calculations of Ψ. The obvious conclusion is that by having a very fine grid in the immediate vicinity of the point (s_α, z_α) and then gradually allowing for a coarser grid further away from that point, we would still get a good solution.

We can compare our results with those in the literature using other methods. In [19], different multidimensional PDE-methods are used than the ones in this thesis (i.e. the Finite Element Method with and without streamline diffusion stabilization), and it is therefore a good source for comparison. In Tables 1 and 2 we present some basic results to compare these different approaches.

		Option Value		
Grid size	Δt	K=95	K=100	K=105
2853	0.01	6.113	1.829	0.163
10226	0.005	6.118	1.848	0.152
38634	0.0025	6.119	1.852	0.150

Table 1: Asian call option (defined in Section 1) values as computed by Zvan et al. [19] ($t = 0$). Here, $r = 0.1, \sigma = 0.1, T = 0.25, S_0 = 100$. Grid size refers to the number of nodes used in the mesh.

		Option Value		
Grid size	Δt	K=95	K=100	K=105
2753	0.01	6.303	3.382	1.255
9858	0.005	5.812	2.232	0.367
38198	0.0025	6.001	2.061	0.136

Table 2: A first result for Asian call option values as computed by the methods of this paper ($t = 0$). Again, $r = 0.1, \sigma = 0.1, T = 0.25, S_0 = 100$.

It should be stated that we use a different mesh than in [19], although with approximately the same number of nodes. Even so, this preliminary result is somewhat disappointing. However, the approach is not as bad as one would think after comparing Tables 1 and 2. The approach does a lot better for larger T and for a larger number of time steps than used in the above example. In Table 3 we have used a mesh with 9858 nodes as in the

σ	K	Broman	Foufas	Večeř	Zvan	MC	L.B	U.B.
0.05	95	10.959	11.112	11.112	11.094	11.094	11.094	11.114
	100	6.679	6.810	6.810	6.793	6.795	6.794	6.810
	105	2.995	2.754	2.750	2.748	2.745	2.744	2.761
0.10	90	15.360	15.416	15.416	15.399	15.399	15.399	15.445
	100	6.989	7.042	7.036	7.030	7.028	7.028	7.066
	110	1.776	1.422	1.421	1.410	1.418	1.413	1.451
0.20	90	15.632	15.659	15.659	15.643	15.642	15.641	15.748
	100	8.419	8.427	8.424	8.409	8.409	8.408	8.515
	110	3.704	3.570	3.568	3.554	3.556	3.554	3.661
0.30	90	16.522	16.533	16.533	16.514	16.516	16.512	16.732
	100	10.228	10.231	10.230	10.210	10.210	10.208	10.429
	110	5.821	5.750	5.748	5.729	5.731	5.728	5.948

Table 3: Asian call option values as computed in this thesis (Broman), Foufas [10], Večeř [17] and Zvan [20] ($t = 0$). The upper and lower bounds (L.B and U.B.) are due to Rogers and Shi [13], and the Monte Carlo (MC) simulations are due to Večeř [17]. The parameters are set to $r = 0.15, T = 1, S_0 = 100$, and for these calculations $\Delta t = 0.01$.

previous example, but used $\Delta t = 0.01$. Even though it is an improvement, this could hardly be described as a very good result. To make a considerable step in the right direction, we thought of two possible approaches. The first and simplest idea was to try to find a reformulation of the original equation, a formulation more suitable for the numerics. The other idea was to try to compensate for the lack of diffusion in the z direction. By studying the resulting solution-surface of a calculation (see below right frame of Figure 1), we can see that this is ill-behaved. The curve "wobbles" a lot, especially for large values of s and z, and this is due to the non-diffusive property of the equation (see [7] for a discussion on hyperbolic equations and hyperbolic properties of equations). Somehow, we want to compensate for this. The theory and results of the first approach is presented in Section 9, but first we take a look at the stabilizing method, referred to as the Streamline Diffusion Method.

8 Streamline Diffusion

In this section we briefly describe the idea behind the streamline-diffusion method. For further reading see [12] or [7]. Equations like the one in this thesis, with diffusion in only one of the two spatial dimensions, often lead to certain numerical instability problems. In the spatial dimension lacking diffusion, the equation becomes hyperbolic in essence, with the associated difficulties. One approach to resolve this problem is to add a small diffusion term in the appropriate spatial direction. Ideally we would like to add the extra diffusion without introducing any errors as the mesh size tends to zero, e.g. we would like to add an equality to the equation. Recall equation (15),

$$\mathcal{L}v = -v_t' - \nabla \cdot (a(x)\nabla v) + b(x) \cdot \nabla v + c(x)v = 0 \tag{20}$$

where

$$a(x) = [\begin{array}{cc} \frac{\sigma^2 s^2}{2} & 0 \\ 0 & 0 \end{array}], \quad b(x) = -[(r-\sigma^2)s\ s^p], \text{ and } c(x) = r.$$

The streamline-diffusion approximation to this problem is similar to the standard finite element approximation, but with one crucial difference, the term with the sum in the following problem formulation. Find $v \in W^1$ such that

$$\begin{aligned} &B_\delta(v,\varphi) \\ &\quad = \int_0^T ((-v_t', \varphi) + (a\nabla v, \nabla\varphi) + (\nabla v, b\varphi) + (cv, \varphi) \\ &\quad + \sum_{\kappa \in \mathcal{T}} \delta_\kappa (\mathcal{L}'v, b \cdot \nabla\varphi)_\kappa) dt = 0 \ \ \forall\, \varphi \in W^0. \end{aligned} \tag{21}$$

On each element of the triangulation, we have defined

$$\mathcal{L}'v = -v_t' - \nabla \cdot (P_\kappa(a\nabla v)) + b \cdot \nabla v + cv$$

where $(\cdot,\cdot)$ and $(\cdot,\cdot)_\kappa$ denotes the inner L_2-product on Ω and κ respectively, and P_κ signifies the orthogonal projection in $L_2(\kappa)$ onto $\mathcal{P}_k$.

The approach is very similar to that of using $\varphi + b \cdot \nabla\varphi$ as test function instead of the usual φ (see equation (17)). However, this is not the case in the streamline-diffusion method (commonly referred to as the SD-method) where $b \cdot \nabla\varphi$ is instead used as a local test function on each element of the triangulation $\mathcal{T}$. We would like to add local equalities to every element in the triangulation $\mathcal{T}$, hence introducing no terms involving the boundaries of the elements. Now, the term of interest in the SD-approach is

$$\sum_{\kappa \in \mathcal{T}} \delta_\kappa (\mathcal{L}'v, b \cdot \nabla\varphi)_\kappa$$

which contains the term

$$\sum_{\kappa \in \mathcal{T}} \delta_\kappa (s^p v_z^{'}, s^p \varphi_z^{'})_\kappa .$$

This term has the same form as a diffusion-term in the z spatial-direction would have in the standard FEM-formulation if it had been present in the original equation. So, we have introduced diffusion by adding

$$\sum_{\kappa \in \mathcal{T}} \delta_\kappa (\mathcal{L}' v, b \cdot \nabla \varphi)_\kappa = \sum_{\kappa \in \mathcal{T}} \delta_\kappa (\nabla \cdot (a \nabla v) - \nabla \cdot (P_\kappa (a \nabla v)), b \cdot \nabla \varphi)_\kappa$$

to equation (17). In the case were a is a constant matrix, this will become zero. In the general case, we hope that the error we are introducing by not solving the original equation any more will be small, and indeed smaller than using the crude approach of simply adding diffusion as mentioned in the beginning of this section (see [14]).

An error analysis of this method can be found in [5]. In the same way as in Section 5, the following matrix equation can be derived by using $\varphi_i = t^0 \phi_i = \phi_i$ as test functions, where $v_n(x)$ is expanded in the nodal basis

$$v_n(x) = \sum_{i=0}^{M(\mathcal{T})} V_{n,i} \phi_i(x).$$

Inserting this into equation (21) yields (with $V_n = (V_{n,0}, V_{n,1}, \ldots, V_{n,M(\mathcal{T})})$

$$\begin{aligned} -(V_n - V_{n-1}) M_1 &+ \frac{k_n}{2}(V_n + V_{n-1}) S_1 + \frac{k_n}{2}(V_n + V_{n-1}) S_2 \\ &+ r(V_n + V_{n-1}) M \frac{k_n}{2} - (V_n - V_{n-1}) S_3 - \frac{k_n}{2}(V_n + V_{n-1}) S_4 \\ &+ \frac{k_n}{2}(V_n + V_{n-1}) S_5 + r \frac{k_n}{2}(V_n + V_{n-1}) S_6 = 0, \end{aligned}$$

where

$$\begin{aligned} M_1 &= (\phi_j, \phi_i), \quad S_1 = (b \cdot \nabla \phi_j, \phi_i), \quad S_2 = (a \nabla \phi_j, \nabla \phi_i), \\ S_3 &= \sum_{\kappa \in \mathcal{T}} \delta_\kappa (\phi_j, b \cdot \nabla \phi_i)_\kappa, \quad S_4 = \sum_{\kappa \in \mathcal{T}} \delta_\kappa (\nabla \cdot (P_\kappa (a \nabla v)) \phi_j), b \cdot \nabla \phi_i)_\kappa, \\ S_5 &= \sum_{\kappa \in \mathcal{T}} \delta_\kappa (b \cdot \nabla \phi_j, b \cdot \nabla \phi_i)_\kappa, \quad S_6 = \sum_{\kappa \in \mathcal{T}} \delta_\kappa (\phi_j, b \cdot \nabla \phi_i)_\kappa . \end{aligned}$$

We look upon this as a system of linear equations that we can easily solve for each time step. Our final condition will determine the last $V_n = V_N$ for us, and we can then successively determine the value of the option for $t = t_n$, until we reach present date, $t = 0$.

As can easily be observed by looking at Figure 3, the SD-method definitely smoothes the solution-curve. But, much to our disappointment the

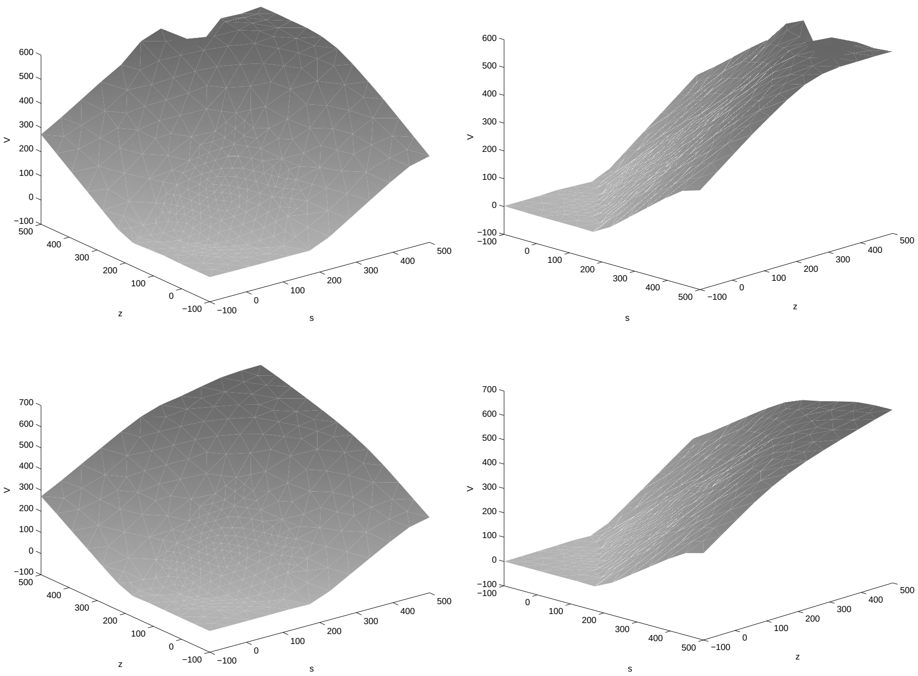

Figure 3: The two pictures at the top describes the solution curve to equation (14) without the use of streamline diffusion. This should be compared to the solution with the use of streamline diffusion to the same problem presented in the pictures at the bottom. Obviously, the curve is smoothened and shows a more "financially feasible" behavior

solution in the point of interest does not improve much! This however is not very surprising since the solution of the dual shows that the influential area and the areas that are stabilized by the SD-method are more or less disjoint! Some (disappointing) results are presented in Table 4. The results for $T = 0.25$ follows the same pattern, no substantial improvement (if indeed any) can be observed. The results are presented in Table 5, and should be compared with Tables 1 and 2.

σ	K	Broman	Foufas	Vecer	Zvan	MC	L.B.	U.B
	95	10.962	11.112	11.112	11.094	11.094	11.094	11.114
0.05	100	6.633	6.810	6.810	6.793	6.795	6.794	6.810
	105	2.966	2.754	2.750	2.748	2.745	2.744	2.761
	90	15.398	15.416	15.416	15.399	15.399	15.399	15.445
0.10	100	7.042	7.042	7.036	7.030	7.028	7.028	7.066
	110	1.480	1.422	1.421	1.410	1.418	1.413	1.451
	90	15.623	15.659	15.659	15.643	15.642	15.641	15.748
0.20	100	8.476	8.427	8.424	8.409	8.409	8.408	8.515
	110	3.529	3.570	3.568	3.554	3.556	3.554	3.661
	90	16.476	16.533	16.533	16.514	16.516	16.512	16.732
0.30	100	10.265	10.231	10.230	10.210	10.210	10.208	10.429
	110	5.712	5.750	5.748	5.729	5.731	5.728	5.948

Table 4: Asian call option values as computed in this thesis (Broman), Foufas [10], Večeř [17] and Zvan [20]. The upper and lower bounds (L.B. and U.B.) are due to Rogers and Shi [13], and the Monte Carlo (MC) simulations are due to Večeř [17]. The parameters are set to $r = 0.15, T = 1, S_0 = 100$, and for these calculations $\Delta t = 0.01$.

		Option Value		
Grid size	Δt	K=95	K=100	K=105
2396	0.01	6.043	2.878	0.898
8495	0.005	5.744	2.194	0.424
33079	0.0025	5.990	2.053	0.211

Table 5: Asian call option values as computed by the Streamline Diffusion method ($t = 0$). Here, $r = 0.1, \sigma = 0.1, T = 0.25, S_0 = 100$. Grid size refers to the number of nodes used in the mesh.

9 Alternative formulation (Derivation of equation, Variational formulation and the Dual problem)

9.1 Derivation of an alternative equation

The first part of this section is devoted to finding an alternative formulation of equation (14). It will follow much the same pattern as in Section 3. Define

$$A(t) = \frac{Z(t)}{t} = \frac{\int_0^t g(S(\lambda), \lambda) d\lambda}{t}$$

called the *running average* of the function $g(S(t), t)$. It follows that

$$dA(t) = (\frac{g(S(t), t)}{t} - \frac{Z(t)}{t^2})dt = (\frac{g(S(t), t)}{t} - \frac{A(t)}{t})dt.$$

We are interested in payoff functions of the form

$$f(S(T), A(T)),$$

so again it seems reasonable that

$$v(t) = v(t, S(t), A(t)). \tag{22}$$

We recall that (equation (4))

$$dS(t) = S(t)(rdt + \sigma d\tilde{W}(t)).$$

Putting

$$X(t) = (S(t), A(t))$$

and applying Itô's lemma to equation (22), results in

$$\begin{aligned} dv(t, X(t)) &= v_t'(t, X(t))dt + v_s'(t, X(t))dS(t) + v_a'(t, X(t))dA(t) \\ &+\frac{1}{2}(v_{ss}''(t, X(t))(dS(t))^2 + 2v_{sa}''(t, X(t))dS(t)dA(t) + v_{aa}''(t, X(t))(dA(t))^2) \\ &= v_t'(t, X(t))dt + v_s'(t, X(t))dS(t) + v_a'(t, X(t))(\frac{g(S(t), t)}{t} - \frac{A(t)}{t})dt \\ &+\frac{S^2(t)\sigma^2}{2} v_{ss}''(t, X(t))dt. \end{aligned}$$

Again, analogously to Section 3 we identify

$$h_s(t) = v_s'(t, X(t))$$

and conclude that

$$h_b(t)rB(t) = v_t'(t, X(t)) + v_a'(t, X(t))(\frac{g(S(t), t)}{t} - \frac{A(t)}{t}) + \frac{S^2(t)\sigma^2}{2} v_{ss}''(t, X(t)).$$

So from equation (5) it follows that

$$v(t,X(t)) = S(t)v_s'(t,X(t)) + \frac{1}{r}(v_t'(t,X(t))$$
$$+v_a'(t,X(t))(\frac{g(S(t),t)}{t} - \frac{A(t)}{t}) + \frac{S^2(t)\sigma^2}{2}v_{ss}''(t,X(t)))$$

from which we conclude that

$$v_t'(t,X(t)) + rS(t)v_s'(t,X(t)) + (\frac{g(S(t),t)}{t} - \frac{A(t)}{t})v_a'(t,X(t))$$
$$+\frac{S(t)^2\sigma^2}{2}v_{ss}''(t,X(t)) - rv(t,X(t)) = 0$$

and with the final condition

$$v(T,S(T),A(T)) = f(T,S(T),A(T)).$$

In the case of our p-mean options

$$v(T,S(T),A(T)) = \max((\frac{Z(T)}{T})^{1/p} - K, 0) = \max(A(T)^{1/p} - K, 0).$$

We are thus led to the equation

$$v_t' + rsv_s' + (\frac{s^p}{t} - \frac{a}{t})v_a' + \frac{s^2\sigma^2}{2}v_{ss}'' - rv = 0, \;\; t > 0, \; s > 0, \; a > 0, \qquad (23)$$

with the final condition

$$v(T,s,a) = \max(a^{1/p} - K, 0).$$

This introduces a singularity in the coefficient of v_a' for $t = 0$. However, since

$$\lim_{t\to 0} \frac{S^p(t) - A(t)}{t} = 0,$$

we will set the coefficient in front of v_a' that is

$$\frac{s^p - a}{t},$$

equal to zero for the last time step. This is a necessary assumption in order to implement the numerics (equation (23) will then simply become the Black-Scholes equation at $t = 0$), and is used for example in [20].

9.2 Variational formulation

Our next step will be to derive the weak formulation of equation (23). Reformulate this equation as

$$-v_t' - \nabla \cdot (\epsilon(x)\nabla v) + b(x) \cdot \nabla v + c(x)v = 0, \quad x = (s, a),$$

where

$$\epsilon(x) = [\begin{matrix} \frac{\sigma^2 s^2}{2} & 0 \\ 0 & 0 \end{matrix}], \quad b(x) = -[(r - \sigma^2)s \; \frac{s^p - a}{t}] \text{ and } c(x) = r.$$

As before, integrate (23) over time and over a suitable bounded area of s and a ($a = A(t)$) denoted by Ω. Using the Gauss divergence theorem gives us

$$\begin{aligned} &\int_0^T \int_\Omega (-v_t' - \nabla \cdot (\epsilon(x)\nabla v) + b(x) \cdot \nabla v + c(x)v)\varphi d\Omega dt \\ &\quad = \int_0^T \int_\Omega (-v_t'\varphi + (\epsilon(x)\nabla v) \cdot \nabla\varphi + (b(x) \cdot \nabla v)\varphi + c(x)v\varphi)d\Omega dt \\ &\quad - \int_0^T \int_\Gamma \varphi(\epsilon\nabla v) \cdot \vec{n} d\Gamma = 0 \end{aligned} \tag{24}$$

as before. Again we seek the weak solution $v \in W$, where

$$W = L^2([0, T]) \times H_1(\Omega), \quad H_1(\Omega) = \{v : \int_\Omega (v^2 + |\nabla v|^2) d\Omega < \infty)\}$$

as the function v such that

$$\begin{aligned} &\int_0^T \int_\Omega (-v_t'\varphi + (\epsilon(x)\nabla v) \cdot \nabla\varphi + (b(x) \cdot \nabla v)\varphi + c(x)v\varphi)d\Omega dt \\ &\quad - \int_0^T \int_\Gamma \varphi(\epsilon\nabla v) \cdot \vec{n} d\Gamma = 0, \quad \forall\varphi \in W. \end{aligned}$$

Next we discretise in the usual way. Let $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$ and $I_n = [t_{n-1}, t_n], k_n = t_n - t_{n-1}$. For each I_n, define $\psi_{n-1}(t) = \frac{t_n - t}{k_n}$ and $\psi_n(t) = \frac{t - t_{n-1}}{k_n}$ for $t \in I_n$. We triangulate Ω into $\mathcal{T} = \{\kappa\}$. As before $\mathcal{P}_k \subset H_1(\Omega)$ denotes the space of all polynomials of degree k defined on each triangle κ of Ω.

Our finite element approximation to (23) will now be to find $v \in W^1$ such that

$$\begin{aligned} &\int_0^T \int_\Omega (-v_t'\varphi + (\epsilon(x)\nabla v) \cdot \nabla\varphi + (b(x) \cdot \nabla v)\varphi + c(x)v\varphi)d\Omega dt \\ &\quad - \int_0^T \int_\Gamma \varphi(\epsilon\nabla v) \cdot \vec{n} d\Gamma = 0 \;\; \forall\varphi \in W^0, \end{aligned} \tag{25}$$

(see Section 5, for the definitions of W^0, and W^1).

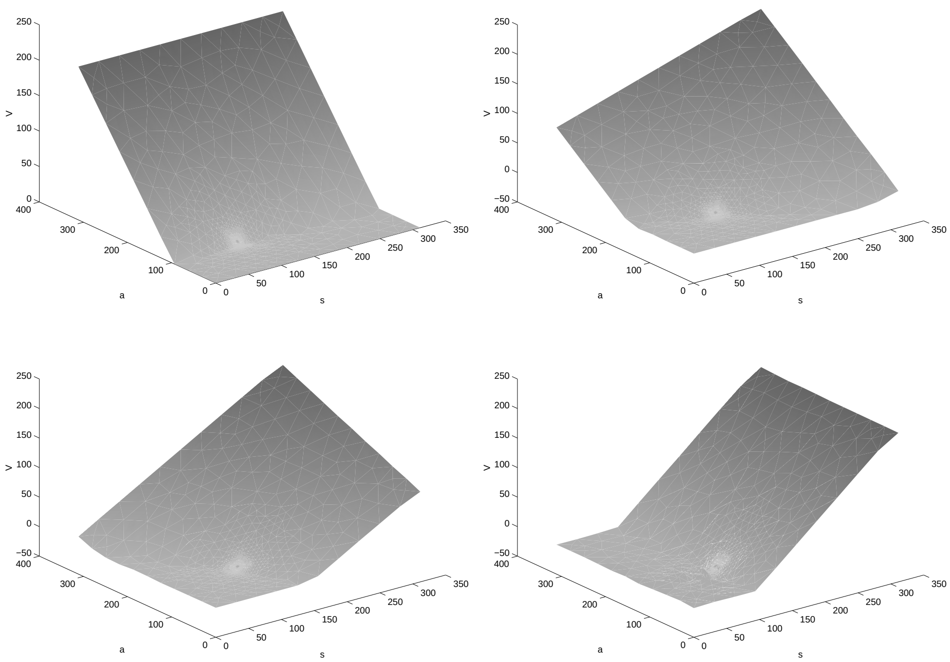

Figure 4: Above left shows $V(t = 0)$, above right $V(t = T/3)$, below left $V(t = 2T/3)$, and below right shows $V(t = T)$.

By the same argument as in Section 5, the last term of equation (25), is set to zero. Expanding $v_n(x)$ in the nodal basis $\{\phi_i\}$ and using ϕ_i as test functions, will lead to the matrix equations

$$V_{n-1}(M(1 + r\frac{k_n}{2}) + \frac{k_n}{2}S_1 + \frac{k_n}{2}S_2) = V_n(M(1 - r\frac{k_n}{2}) - \frac{k_n}{2}S_1 - \frac{k_n}{2}S_2),$$

where,

$$M = (\phi_j, \phi_i), \quad S_1 = (b(s,a) \cdot \nabla\phi_j, \phi_i), \quad S_2 = (\epsilon(s,z)\nabla\phi_j, \nabla\phi_i).$$

just as before. The difference lies in the definition of $b(s,a)$. A set of solution curves are presented in Figure 4.

9.3 The Dual

Again, by considering the adjoint operator $\mathcal{L}^*$, defined by $\langle \mathcal{L}v, \Psi \rangle = \langle v, \mathcal{L}^*\Psi \rangle$, ($\langle \cdot, \cdot \rangle$ denoting $\int_0^T \int_\Omega$) we can formulate the dual problem to equation (23). Recalling that

$$\mathcal{L} = \frac{\partial}{\partial t} + rs\frac{\partial}{\partial s} + \frac{s^p - a}{t}\frac{\partial}{\partial a} + \frac{s^2\sigma^2}{2}\frac{\partial^2}{\partial s^2} - r,$$

it is easy to conclude simply by redoing the calculations leading to $\langle v, \mathcal{L}^*\Psi \rangle$ from $\langle \mathcal{L}v, \Psi \rangle$, that the dual problem must be

$$\begin{aligned} \mathcal{L}^*\Psi = -\Psi_t' - (2r - \frac{1}{t} - \sigma^2)\Psi + (\sigma^2 - r)s\Psi_s' & \\ -\frac{s^p - a}{t}\Psi_a' + \frac{\partial}{\partial s}(\frac{\sigma^2 s^2}{2}\Psi_s') &= 0 \\ \Psi(t = 0, s, a) &= \delta_{s_\alpha, a_\alpha} \\ \Psi(t, s, a) &= 0, \quad t > 0, \ (s, a) \in \Gamma. \end{aligned} \tag{26}$$

Now, to see why this is relevant, we multiply the equation above by a function e, and transfer the derivatives of Ψ, onto e

$$\begin{aligned} &\int_0^T \int_\Omega (-\Psi_t' - (2r - \frac{1}{t} - \sigma^2)\Psi + (\sigma^2 - r)s\Psi_s' \\ &\quad -\frac{s^p - a}{t}\Psi_a' + \frac{\partial}{\partial s}(\frac{\sigma^2 s^2}{2}\Psi_s'))e \, dtd\Omega \\ &= \int_\Omega (\Psi(0)e(0) - \Psi(T)e(T))d\Omega \\ &\quad + \int_0^T \int_\Omega (\Psi e_t' - (\sigma^2 - r)\Psi\frac{\partial}{\partial s}(se) + \Psi\frac{\partial}{\partial a}(\frac{s^p - a}{t}e) \\ &\quad -(\frac{\sigma^2 s^2}{2}\Psi_s')\frac{\partial}{\partial s}(e) - (2r - \frac{1}{t} - \sigma^2)\Psi e)dtd\Omega \\ &= \int_\Omega (\Psi(0)e(0) - \Psi(T)e(T))d\Omega \\ &\quad + \int_0^T \int_\Omega \Psi(e_t' + rse_s' + \frac{s^p - a}{t}e_a' + \frac{s^2\sigma^2}{2}e_{ss}'' - re)dtd\Omega \\ &= \int_\Omega (\Psi(0)e(0) - \Psi(T)e(T))d\Omega + \int_0^T \int_\Omega (\Psi\mathcal{L}e)dtd\Omega = 0. \end{aligned}$$

Choose $e = v - V$, where v solves equation (23), and V is the approximate solution, i.e., the solution to equation (25). Since $v(T)$ is known, $e(T) = 0$, and this in combination with $\Psi(t = 0, s, a) = \delta_{s_\alpha, a_\alpha}$, makes us conclude that

$$e(t = 0, s_\alpha, a_\alpha) = -\int_0^T \int_\Omega (\Psi\mathcal{L}(v - V))dtd\Omega = \int_0^T \int_\Omega (\Psi\mathcal{L}V)dtd\Omega. \tag{27}$$

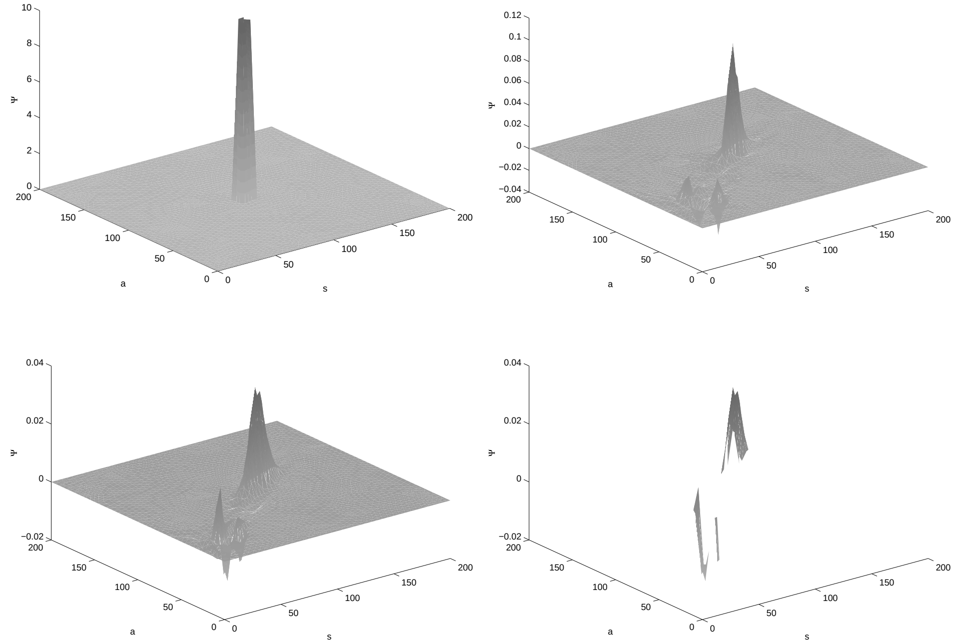

Figure 5: Above left, is showing $\Psi(t = 0)$, above right $\Psi(t = T/2)$, below left shows $\Psi(t = T)$, and below right shows $\Psi(t = T)$ as well, but only values of Ψ such that $\Psi > \frac{1}{100}$.

Again, we are interested in the behavior of Ψ, and some numerical results are presented in Figure 5. To illuminate the outcome, it is also useful to look at the result from above, and focus color-intensity to regions where Ψ is large (see Figure 6).

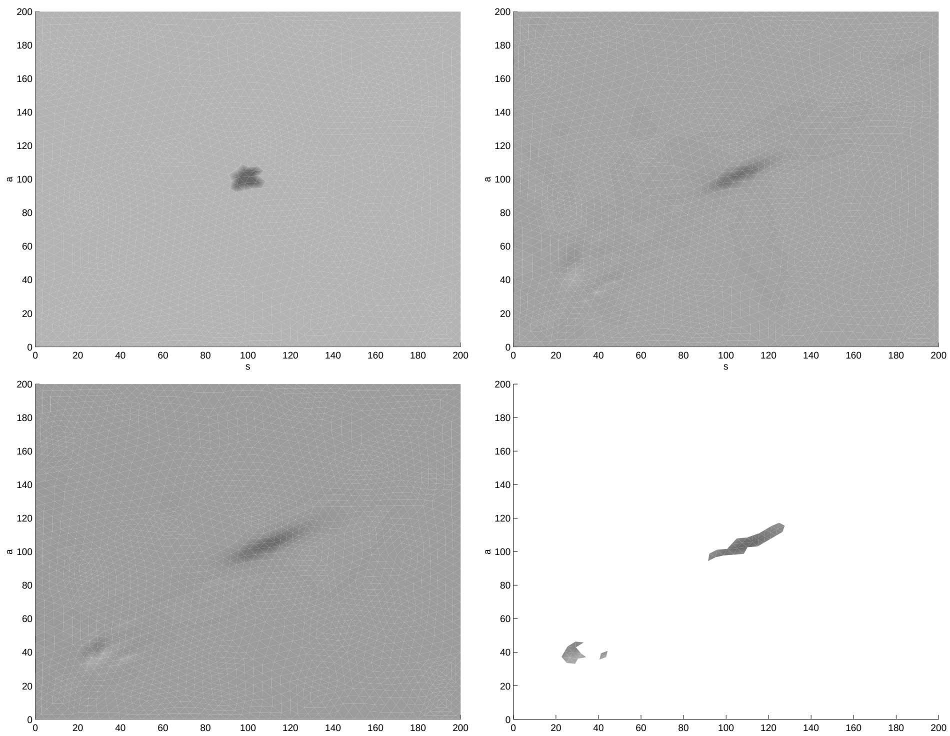

Figure 6: Above left, is showing $\Psi(t = 0)$, above right $\Psi(t = T/2)$, below left shows $\Psi(t = T)$, and below right shows $\Psi(t = T)$ as well, but only values of Ψ such that $\Psi > \frac{1}{100}$.

10 Further results

In Table 6, we present some results as obtained with the alternative formulation of the governing equation.

		Option Value		
Grid size	Δt	K=95	K=100	K=105
2279	0.01	6.114	1.849	0.156
7546	0.005	6.118	1.850	0.149
32897	0.0025	6.119	1.851	0.149

Table 6: Asian call option values as computed in this thesis ($t = 0$). Again, $r = 0.1, \sigma = 0.1, T = 0.25, S_0 = 100$.

This result should be compared with the results of Tables 1 and 2. It is not hard to see that the results obtained here are far better than those obtained from the original formulation. Furthermore, the results here compared to the results of Zvan et al. [19] are more exact even though fewer nodes are used. To demonstrate this even further, Table 3 is reproduced and presented in Table 7 although using equation (23) instead of equation (14) for calculating our solution. The column representing the calculations of this paper, is very close to the results of Zvan et al. and the results of the Monte-Carlo method. These are also very close to the lower bound, and as is pointed out in [13] this is in fact very close to the true value, much closer than the upper bound. In Tables 8 and 9, we present the results from a number of different calculations using different mesh sizes and different numbers of time steps for the same setting of parameters. For comparison, the calculations made in this thesis from Table 7 are considered to be the "exact" solution. Since there is no (known) exact solution, this numerical solution will be considered a reference solution. It is easy to see that relatively few nodes and time steps need to be used to obtain a reasonably good result. It should be remembered that for practical purposes a retailer would be satisfied with an accuracy on the percentage level. Already the mesh used in Table 8 could suffice, even though it performs badly especially in the case of $\sigma = 0.1,\ K = 110$. The setting in Table 9 is satisfactory, but by taking something in-between the two cases we could find an acceptably accurate but still computationally efficient parameter setting.

σ	K	Broman	Foufas	Vecer	Zvan	MC	L.B.	U.B
	95	11.094	11.112	11.112	11.094	11.094	11.094	11.114
0.05	100	6.795	6.810	6.810	6.793	6.795	6.794	6.810
	105	2.746	2.754	2.750	2.748	2.745	2.744	2.761
	90	15.399	15.416	15.416	15.399	15.399	15.399	15.445
0.10	100	7.028	7.042	7.036	7.030	7.028	7.028	7.066
	110	1.420	1.422	1.421	1.410	1.418	1.413	1.451
	90	15.642	15.659	15.659	15.643	15.642	15.641	15.748
0.20	100	8.408	8.427	8.424	8.409	8.409	8.408	8.515
	110	3.556	3.570	3.568	3.554	3.556	3.554	3.661
	90	16.512	16.533	16.533	16.514	16.516	16.512	16.732
0.30	100	10.208	10.231	10.230	10.210	10.210	10.208	10.429
	110	5.728	5.750	5.748	5.729	5.731	5.728	5.948

Table 7: Asian call option values as computed in this thesis (Broman), Foufas [10], Večeř [17] and Zvan [20] ($t = 0$). The upper and lower bounds (L.B. and U.B.) are due to Rogers and Shi [13], and the Monte Carlo (MC) simulations are due to Večeř [17]. The parameters are set to $r = 0.15, T = 1, S_0 = 100$, and for these calculations $\Delta t = 0.01$. The number of nodes are 41869.

σ	K	Exact	This run	R.E. in %
	95	11.094	11.088	0.054
0.05	100	6.795	6.788	0.103
	105	2.746	2.759	0.473
	90	15.399	15.396	0.019
0.10	100	7.028	7.026	0.028
	110	1.420	1.376	3.099
	90	15.642	15.639	0.019
0.20	100	8.408	8.405	0.036
	110	3.556	3.525	0.872
	90	16.512	16.502	0.061
0.30	100	10.208	10.198	0.098
	110	5.728	5.702	0.279

Table 8: Asian call option ($t = 0$). The mesh consists of 2469 nodes, the number of time steps is 25.

σ	K	Exact	This run	R.E. in %
	95	11.094	11.093	0.009
0.05	100	6.795	6.791	0.059
	105	2.746	2.745	0.036
	90	15.399	15.397	0.013
0.10	100	7.028	7.022	0.083
	110	1.420	1.419	0.070
	90	15.642	15.639	0.019
0.20	100	8.408	8.402	0.071
	110	3.556	3.560	0.112
	90	16.512	16.509	0.018
0.30	100	10.208	10.200	0.078
	110	5.728	5.730	0.035

Table 9: Asian call option. The mesh consists of 8313 nodes, the number of time steps is 60.

11 General p-mean options

To have something to compare the results with from the calculations using FEM, a simple Monte-Carlo program was written. In Tables 10 and 11, we present some results from calculations with different p using this simple Monte-Carlo program. We used two million simulations of the stock price trajectory, and as can be seen by comparing the first column of Table 10 to the appropriate column of Table 9, the results are within reasonable error limits. Table 11 should be compared to the results of Tables 1,2 and 6. There are two major sources for errors. The first is the limited number of stock price trajectory simulations and the other is that a trajectory is simulated by a finite (in this case 1000) rather than infinite number of steps.

The general payoff for a p-mean option has the form

$$f(T, S(T), A(T)) = \max((\int_0^T \frac{1}{T} S^p(\lambda) d\lambda)^{1/p} - K, 0).$$

By letting $p \to \infty$, this will lead us to the payoff of a lookback option, i.e. an option with the payoff

$$f(T, S(T)) = \max(\max_{0 \leq \lambda \leq T} S(\lambda) - K, 0).$$

Unfortunately, the form of equation (14) is very impractical for the Finite Element approach for $p > 1$. The relevant solution area has to have the proportions

$$\Omega = [0, s_{max}] \times [0, s_{max}^p] = [0, s_{max}] \times [0, a_{max}],$$

K	p=1	p=2	p=5	p=700	lb (MC)	lb (exact)
90	15.666	16.018	17.122	31.296	32.284	32.702
100	8.393	8.765	9.717	22.692	23.660	24.095
110	3.549	3.865	4.614	15.100	17.015	17.818

Table 10: General p-mean options, here we have used r=0.15, $\sigma = 0.2$, T=1 and $s_0 = 100$. The notation lb stands for lookback and the exact price of this option is calculated using the formula on page 59 of [21].

Option Value		
K=95	K=100	K=105
6.112	1.853	0.149

Table 11: Asian call option values as computed in this thesis ($t = 0$). Again, $r = 0.1, \sigma = 0.1, T = 0.25, S_0 = 100$ and $p = 1$.

since if $s = s_{max}$ $\forall t \in [0, T] \Rightarrow a = s^p_{max}$. For p large, Ω will become very thin in one spatial dimension compared to the other. One approach to get around this would be to use the scaling property of equation (23). Let

$$S(t) \to cS(t) \text{ and } \quad K \to cK$$

which implies that

$$A(t) \to c^p A(t) \text{ and } f(T) \to cf(T).$$

By letting $c = \frac{1}{s_{max}}$ and solving for $u = cv$, we can solve the equation in the domain

$$\Omega = [0, cs_{max}] \times [0, (cs_{max})^p] = [0, cs_{max}] \times [0, c^p a_{max}] = [0, 1] \times [0, 1],$$

but retrieve the desired solution value $v(0)$ by multiplying u by $1/c$; $v(0) = u(0)/c$. This makes it possible to go up to at least $p = 3$ with fairly good results (see Table 12 and compare with Table 13).

K	p=2	p=3	p=5
90	16.012	16.358	55.689
100	8.732	9.233	52.670
110	3.680	3.910	49.651

Table 12: General p-mean options calculated using FEM, we have used r=0.15, $\sigma = 0.2$, T=1 and $s_0 = 100$.

K	p=2	p=3	p=5	p=10
90	16.018	16.389	17.122	18.745
100	8.765	9.086	9.717	11.200
110	3.865	4.091	4.614	5.797

Table 13: General p-mean options using Monte Carlo methods, we have used r=0.15, $\sigma = 0.2$, T=1 and $s_0 = 100$.

Another difficulty arises however, as is obvious by studying the results for $p = 5$. We have that $A(0) = S^p(0)$, and since we have transformed $S(0)$ to be less than one, $A(0)$ will become very close to zero. It would be necessary to have an extremely fine grid to be able to resolve the values of $A(0) = a_0$. Furthermore, it is very close to the boundary, and the earlier argument of not requiring an accurate boundary condition is not valid anymore. Again we will re-write equation (14). This time, choose

$$A(t) = (\frac{1}{t}\int_0^t S^p(\lambda)d\lambda)^{1/p}$$

as our second spatial variable. We conclude that

$$\begin{aligned}
dA(t) &= \left(-\frac{1}{p}t^{-(1/p+1)}(\int_0^t S^p(\lambda)d\lambda)^{1/p} + t^{-1/p}\frac{S^p(t)}{p}(\int_0^t S^p(\lambda)d\lambda)^{1/p-1}\right)dt \\
&= \left(-\frac{1}{pt}A(t) + \frac{S^p(t)}{pt}(\frac{1}{t}\int_0^t S^p(\lambda)d\lambda)^{1/p-1})\right)dt \\
&= \left(-\frac{1}{pt}A(t) + \frac{S^p(t)}{pt}A^{1-p}(t)\right)dt = \left(\frac{A^{1-p}(t)}{pt}(S^p(t) - A^p(t))\right)dt.
\end{aligned}$$

Once again, the assumption

$$v(t) = v(t, S(t), A(t)),$$

and Itô's lemma yields

$$\begin{aligned}
dv(t, X(t)) &= v_t'(t, X(t))dt + v_s'(t, X(t))dS(t) + v_a'(t, X(t))dA(t) \\
&+\frac{1}{2}(v_{ss}''(t, X(t))(dS(t))^2 + 2v_{sa}''(t, X(t))dS(t)dA(t) + v_{aa}''(t, X(t))(dA(t))^2) \\
&= v_t'(t, X(t))dt + v_s'(t, X(t))dS(t) + v_a'(t, X(t))\frac{A^{1-p}(t)}{pt}(S^p(t) - A^p(t))dt \\
&+\frac{S^2(t)\sigma^2}{2}v_{ss}''(t, X(t))dt.
\end{aligned}$$

Once more we identify

$$h_s(t) = v_s'(t, X(t))$$

and conclude that

$$h_b(t)rB(t) \\ = v_t'(t,X(t)) + v_a'(t,X(t))\frac{A^{1-p}(t)}{pt}(S^p(t)-A^p(t)) + \frac{S^2(t)\sigma^2}{2}v_{ss}''(t,X(t)).$$

So from equation (5) it follows that

$$v(t,X(t)) = S(t)v_s'(t,X(t)) + \frac{1}{r}(v_t'(t,X(t)) \\ + v_a'(t,X(t))\frac{A^{1-p}(t)}{pt}(S^p(t)-A^p(t)) + \frac{S^2(t)\sigma^2}{2}v_{ss}''(t,X(t))),$$

from which we conclude that

$$v_t'(t,X(t)) + rS(t)v_s'(t,X(t)) + \frac{A^{1-p}(t)}{pt}(S^p(t)-A^p(t))v_a'(t,X(t)) \\ + \frac{S(t)^2\sigma^2}{2}v_{ss}''(t,X(t)) - rv(t,X(t)) = 0.$$

Moreover we have the final condition

$$v(T,S(T),A(T)) = f(T,S(T),A(T)),$$

which now is

$$v(T,S(T),A(T)) = \max(A(T)-K,0)$$

We are thus led to the equation

$$v_t' + rsv_s' + \frac{a^{1-p}}{pt}(s^p-a^p)v_a' + \frac{s^2\sigma^2}{2}v_{ss}'' - rv = 0,\ t>0,\ s>0,\ a>0. \quad (28)$$

with the final condition

$$v(T,s,a) = \max(a-K,0)$$

Again, this introduces a singularity in the coefficient of v_a' for $t=0$, and since

$$\lim_{t\to 0}\frac{S^p(t)-A^p(t)}{t} = 0,$$

we will set

$$\frac{s^p-a^p}{t} = 0,$$

for the last time step. This is a necessary assumption in order to implement the numerics (equation (28) will then simply become the Black-Scholes equation at $t=0$).

With the new formulation $A(0)=S(0)$, and a_0 will not "wander away" towards the boundary as p increases. Of course, it would be too simple if this resolved our difficulties altogether. We can see from equation (28) that a must not equal zero. This is actually a valid assumption, a can never actually equal zero but can come arbitrarily close. We will therefore modify our computational domain, and use

$$\Omega = [0, s_{max}] \times [a_{min}, s_{max}] = [0, s_{max}] \times [a_{min}, a_{max}],$$

before again re-scaling Ω. We justify this in part by the results it produces (see Table 14) and in part by the argument of localized behavior mentioned in Sections 5, and 9.2. Another source of problem is that the coefficient of $v_a^{'}$ will become extremely large for small a:s and large s:s, making equation (28) exhibit hyperbolic behavior. The resolution of this difficulty is left as an open question, suggesting the Streamline Diffusion method as a possible approach.

K	p=2	p=3	p=5	p=10
90	15.979	16.394	17.101	18.800
100	8.714	9.082	9.721	11.243
110	3.801	4.090	4.603	5.803

Table 14: General p-mean options calculated using FEM, we have used r=0.15, $\sigma = 0.2$, T=1 and $s_0 = 100$.

References

[1] L. Bachelier. *Théorie de la spéculation.* Annales scientifiques de l'École Normale Supérieure 17, pp 21-86 1900.

[2] J. Barraquand and T. Pudet. *Pricing of American Path-Dependant Contingent Claims.* Mathematical Finance 6(1), 17-51, 1996.

[3] F. Black, M. Scholes. *The Pricing of Options and Corporate Liabilities.* Journal of Political Economy 81, 637-654, 1973.

[4] C. Borell. *Matematik och Optioner,* lecture notes Chalmers University of Technology, Göteborg, 1999.

[5] E. Broman *A stability analysis of a time dependent Streamline Diffusion Method,* working paper, Department of Mathematics Chalmers University of Technology, Göteborg.

[6] L. Debnath and P. Mikusiński. *An Introduction to Hilbert Spaces with Applications, 2nd edition.* Academic Press, 1999.

[7] K. Eriksson D. Estep P. Hansbo and C. Johnson. *Computational Differential Equations,* Studentlitteratur, Lund, 1996.

[8] G.B. Folland. *Real Analysis. Modern Techniques and Their Applications, 2nd edition,* John Wiley & Sons INC., 1999.

[9] P.A. Forsyth, K.R. Vetzal, and R. Zvan. *Convergence of lattice and PDE methods for pricing Asian Options.* Working paper.

[10] G. Foufas. *PDE-methods for Asian options.* Master's Thesis, Chalmers University of Technology, Göteborg, 2000.

[11] A. Friedman. *Stochastic differential equations and applications.* Academic Press, 1975.

[12] P. Houston and E. Süli. *Stabilized* hp-*Finite Element Approximation of Partial Differential Equations with Nonnegative Characteristic Form.*

[13] L. Rogers and Z. Shi. *The value of an Asian option.* Journal of Applied Probability, 32:1077-1088, 1995.

[14] H.-G. Roos, M. Stynes, and L. Tobiska. *Numerical Methods for Singularly Perturbed Differential Equations. Convection-Diffusion and Flow Problems,* Volume 24 of *Springer Series in Computational Mathematics.* Springer-Verlag, 1996.

[15] P.A. Samuelson. *Rational theory of warrant pricing.* Industrial Management Rev. 6, 13-32 1965.

[16] Ch. Schwab. *p- and hp-Finite Element Methods. Theory and Applications to Solid and Fluid Mechanics.* Oxford University Press, 1998.

[17] J. Večeř. *A new PDE approach for pricing arithmetic average Asian options.* The Journal of Computational Finance, 2000.

[18] P. Wilmott, J.N. Dewynne, and S Howison (1993). *Option Pricing: Mathematical Models and Computation.* Oxford: Oxford Financial Press.

[19] R. Zvan, P.A. Forsyth and K.R. Vetzal. *A Finite Volume Approach For Contingent Claims Valuation* Working paper.

[20] R. Zvan, P.A. Forsyth and K. Vetzal. *Robust Numerical Methods for PDE Models of Asian Options*, The Journal of Computational Finance, 1(2), Winter 1997/98.

[21] A. Öhgren. *On the Pricing of Lookback Options.* Master's Thesis, Chalmers University of Technology, 1999.