APPROXIMATE SOLUTION TO SECOND ORDER PARABOLIC EQUATIONS, WITH APPLICATION TO FINANCIAL MODELING

A Dissertation in Mathematics
by
Chao Liang

© 2014 Chao Liang

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

August 2014
The dissertation of Chao Liang was reviewed and approved* by the following:

Victor Nistor  
Professor of Mathematics  
Dissertation Co-Advisor, Co-Chair of Committee

Xiantao Li  
Associate Professor of Mathematics  
Dissertation Co-Advisor, Co-Chair of Committee

Anna Mazzucato  
Professor of Mathematics

John Liechty  
Professor of Marketing and Statistics

Qiang Du  
Professor of Mathematics

Yuxi Zheng  
Professor of Mathematics  
Head of the Department of Mathematics

*Signatures are on file in the Graduate School.
Abstract

In this dissertation, we consider second order parabolic equations with variable coefficients. We derive the closed-form approximations to the associated fundamental solution, as well as general solutions for certain initial conditions. The starting point is the Dyson-Taylor commutator method which was recently developed. We derive the approximate formula in a new form, which is easy to be calculated by computer programming. We next show that the Dyson series can be computed without Taylor expansion, which leads to more elementary computations, as well as cleaner results. In the section on theory, we have proven that the approximate solution is asymptotic series under certain regularity conditions; in the section for practical application, we have successfully implemented symbolic computation to several popular models in finance: CEV model, SABR model and Heston model. We not only approximate the fundamental solution (transition kernel), but also the option price (which is important for calibration and pricing), the implied volatility (which is quoted by traders), as well as the characteristic function of the underlying stochastic process (which is important for understanding the volatility skew/smile). The computation is carried out by MATLAB codes, and they are available upon request to charlesleung2009@gmail.com or cul192@psu.edu
# Table of Contents

List of Figures vii  
List of Symbols ix  
Acknowledgments x  

Chapter 1  
**Introduction and Literature Review** 1  
1.1 Beyond the Black-Scholes Model 1  
1.2 Analytical/Numerical Approaches to Heat Kernel/Option Pricing/Implied Volatility 3  
1.2.1 Approximation to the transition probability density 6  
1.2.2 Approximation to the option price 9  
1.2.3 Approximation to the implied volatility 10  
1.2.4 Approximation to the characteristic functions 13  
1.2.5 When coefficients are time-dependent 14  
1.3 Organization of the dissertation 15  

Chapter 2  
**Formal Expansion for Small Parameter Perturbation** 18  
2.1 The Commutators 19  
2.2 Application of the Commutators to the Approximation Problem 20  
2.3 An Example: Small Parameter Expansion for the CEV Model 24  

Chapter 3  
**Small Time Expansion for the Heat Kernel-Reformulate the Dyson-Taylor Commutator method** 28  
3.1 Local Dilations and Perturbative Expansions 30  
3.2 Convergence of the short time expansion in Certain function space 36  
3.3 Duhamel’s principle and the Iterative Approach 36
Chapter 4
Approximating the Solutions of Parabolic Equations-Dyson-Taylor and Beyond

4.1 Approximation for the solution .................................................. 41
4.2 Different approaches for the approximation of the heat kernel ........ 44
4.3 A splitting method where Taylor expansion is not needed .......... 46
   4.3.1 An Example: Matrix Case ................................................. 47
   4.3.2 Operator splitting and a new approach to compute matrix exponential 48
   4.3.3 Differential operator case ................................................. 49
   4.3.4 Weighted Sobolev space ............................................... 49
   4.3.5 Proof of asymptotic expression ....................................... 51

Chapter 5
Two and Higher Dimensional Approximation: Heston, SABR models 54

5.1 Higher dimensions: when the initial condition only depends on the first coordinate ............................................. 54
   5.1.1 Short time expansion for the Heston’s model ....................... 56
   5.1.2 Short time expansion for the SABR model ......................... 57

5.2 Transition kernel for the SABR model .................................. 59

Chapter 6
Extending the Approximating Formula to Time/Space Dependent Coefficients 63

6.1 Time/Space dependent coefficients ........................................ 63
6.2 A Example:Hull-White model for the Bond Price ..................... 64

Chapter 7
American Style Options ......................................................... 67

7.1 American put option for the CEV model with zero interest rate ....... 67
7.2 Heat kernel expansion for CEV-first attemp ............................ 69

Chapter 8
Jump-Diffusion Models .......................................................... 72

8.1 Review of Lévy processes .................................................. 73
   8.1.1 The Poisson Process .................................................. 73
   8.1.2 Compound Poisson processes ........................................ 74
   8.1.3 Jump-diffusion and Lévy processes ................................ 74
   8.1.4 Gamma processes .................................................. 75
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1.5</td>
<td>Variance Gamma process</td>
<td>75</td>
</tr>
<tr>
<td>8.1.6</td>
<td>Inverse Gaussian Process</td>
<td>76</td>
</tr>
<tr>
<td>8.1.7</td>
<td>Normal Inverse Gaussian process</td>
<td>76</td>
</tr>
<tr>
<td>8.1.8</td>
<td>Strictly $\alpha$-stable processes</td>
<td>76</td>
</tr>
<tr>
<td>8.2</td>
<td>Exponential Lévy models</td>
<td>77</td>
</tr>
<tr>
<td>8.2.1</td>
<td>A survey of popular jump models</td>
<td>77</td>
</tr>
<tr>
<td>8.3</td>
<td>Outside Fourier Transform regime: commutator method applied to Kou's</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>model with local volatility</td>
<td></td>
</tr>
<tr>
<td>8.4</td>
<td>Local volatility jump-diffusion models with stochastic volatility</td>
<td>81</td>
</tr>
</tbody>
</table>

**Chapter 9**

**Approximate the Implied Volatilities/Heat Kernels for Various Models**  
9.1 Approximation to the implied volatility/heat kernel of CEV model  
9.2 Approximation to the implied volatility/heat kernel of SABR model  
9.3 Approximation to the implied volatility of Heston model-A different approach

**Chapter 10**

**Conclusion and Future Research Directions**

**Appendix A**

**Computing Operators Multiplications and Commutators**  
A.1 Representation of differential operator by matrix and compute their multiplication  
A.1.1 One dimensional operator multiplication and commutator computation  
A.1.2 Two and higher dimensions  
A.2 Compute $P_n(t)$  
A.2.1 Compute $P_n(t)$ inductively  
A.2.2 Efficient way of computing for two and higher dimensions  
A.3 Compute closed form approximation to option price/heat kernel/implied volatility  
A.3.1 Computation of the implied volatility  
A.3.1.1 Compute higher order derivatives - Hermite polynomials  
A.3.2 Algorithm to compute the closed form approximation to option price  
A.3.3 Algorithm to compute the closed form approximation to heat kernel

**Bibliography**
List of Figures

2.1 CEV model for small alpha value ........................................... 27
5.1 Heston model European call price, with \( t = .25, \sigma_v = .32, \theta = .36, \rho = -.01, V_0 = .24, S_0 = 100, r = .05, \) and various strikes ........................................... 57
5.2 \( \lambda \) SABR model European call price, with \( t = .25, \beta = .85, \kappa = 4, \theta = .34, \alpha = .22, \rho = -.04, \sigma_0 = .35, S_0 = 100, r = .05, \) and various strikes ........................................... 59
5.3 Transition kernel for SABR .................................................... 62
6.1 Hull White Model for the Bond Price, with linear dependence of time ........................................... 66
7.1 CEV model European put option price, with \( t = 1.5, \sigma = .34, \alpha = .65, S_0 = 100, r = 0, \) and various strikes ........................................... 68
7.2 Heat kernel expansion for CEV model ........................................ 70
7.3 expansive approximation to CEV model for American put options ........................................... 71
8.1 Kou’s model with constant coefficients ........................................ 82
8.2 Kou’s model for \( \lambda = .1, S_0 = 100, \sigma = .34, \rho = .4, \eta_1 = 2.1, \eta_2 = 1.9, t = .75, r = .05, \) and local volatility \( h(x) = \exp(-.25(x - \log(S_0))) \), compared to Monte Carlo simulation solution for various strikes ........................................... 83
8.3 Bates model European call price, with \( \kappa = 4, \theta = .24, \sigma_v = .23, \rho = -.03, \lambda = .05, \mu_J = .0001, \sigma_J = .01, t = .25, V_0 = .34^2, S_0 = 30, r = .05\), and various strikes ........................................... 84
8.4 Bates model European call price, with \( \kappa = 4, \theta = .24, \sigma_v = .23, \rho = -.03, \lambda = .1, \mu_J = .0001, \sigma_J = .1, t = .25, V_0 = .34^2, S_0 = 30, r = .05, \) with local volatility in the form of \( h(x) = \exp((.7 - 1)(x - \log(S_0))) \), and various strikes ........................................... 85
9.1 CEV model European call price, with \( t = 5.25, \sigma = .34, \alpha = .5, S_0 = 90, r = 0, \) and various strikes ........................................... 89
9.2 SABR model European call price, with \( \beta = .85, \alpha = .22, \rho = -.04, \sigma_0 = .35, t = .75 \) for first figure, and \( t = 7.25 \) for the second, \( S_0 = 100, \) with various strikes ........................................... 93
9.3 Heston model European call price, with $t = .25, \kappa = 5, \theta = .2, \sigma_V = .26, \rho = - .01, V_0 = .14, S_0 = 50, r = .05$, and various strikes .......................... 97

A.1 Code to compute $S$ .......................................................... 103
A.2 Matlab code to compute $P_n(t)$ ................................. 106
A.4 Higher order approximation for implied volatility of CEV model ......... 113
A.5 Higher order approximation for implied volatility of CEV model .......... 113
A.6 Higher order approximation for implied volatility of CEV model ........... 114
List of Symbols

$S$  Current stock price
$F$  Current forward price
$K$  Strike price of an option
$\sigma$  Volatility
$\rho$  Correlation
$T$  Time to maturity
$e^{tL}$  The semigroup generated by a time-independent operator $L$
$\mathbb{L}_\gamma$  Strongly elliptic operator with uniform bound
$ad_L$  Adjoint operator
$\mathcal{G}$  The Green’s function, or transition probability density function
$W^{k,p}(\Omega)$  The standard Sobolev space with indices $k$ and $p$ on the domain $\Omega$
$W^{k,p}_a(\Omega)$  The exponentially weighted Sobolev space with indices $k$ and $p$ and $a$, on the domain $\Omega$
$\mathbb{R}^N$  $N$-dimensional Euclidean space
Acknowledgments

I would like to express my sincere gratitude to my thesis advisors Victor Nistor and Xiantao Li, for their continuous support of my Ph.D. study at Penn State University. They taught me partial differential equations and numerical analysis, which are two main topics of this thesis. They gave me financial support from their own grant, such that I can focus on the research. More importantly, after they showed me the direction on numerical/closed-form approximate solutions to PDE, they also gave me certain level of freedom to choose my area of study, and were supportive when I decided to take an internship in financial industry to get first hand experience, which ultimately lead to this thesis with application to financial modeling.

Besides my advisors, I am also indebted to my committee members. I discussed with Professor John Liechty once every week during the last year, which helped me to learn a lot on statistics and finance, especially on Markov Chain Monte Carlo simulation on calibrating models. I discussed a lot with Professor Anna Muzzucato, especially during the last three months; moreover, she gave me financial support for further research as I was graduating. I also learned a lot from Professor Qiang Du, on numerical PDEs.

Several Professors in Penn State are very helpful in my research: Professor Alberto Bressan, Professor Alexei Novikov, Associate Professor Tao Yao, Assistant Professor Heber Farnsworth.

Last but not the least, I would like to thank my wife Zhan Huang, for her continuous support to our family.
Dedication

To my parents, Da Liang and Guoqun Shu.
Chapter 1  
Introduction and Literature Review

1.1 Beyond the Black-Scholes Model

During the past thirty year, we have seen the fast development of derivatives pricing theory and practice. The Black-Scholes formula \[1\] has been widely used to price European options. In the Black-Scholes setting, assets are assumed to follow constant volatility log-normal processes

\[
dS_t = S_t(\mu dt + \sigma dW_t). \tag{1.1}
\]

Since then, a huge amount of literature was developed to propose new models and analytical/numerical methods to explain the volatility smiles. All these models follows the original idea of Black-Scholes, that the European style option price is the expected payoff discounted in the risk neutral measure \[2\]. In summary, there are three main approaches:

1. The volatility \( \sigma \) in (1.1) is not a constant but rather a deterministic function of the underlying asset and time. These types of models are called local volatility models. For them, the dynamics of the underlying asset is governed by the SDE:

\[
dS_t = S_t(\mu dt + \sigma(S_t, t)dW_t) \tag{1.2}
\]

A remarkable result is due to Dupire \[3\], states that one can get recover unique local volatility by the relation:

\[
\sigma(K, T) = \sqrt{2 \frac{C_T + rKC_K}{K^2C_{KK}}} \tag{1.3}
\]

where \( C_K, C_{KK} \) are the first- and second-order derivatives of the call price with respect to its strike and \( C_T \) the derivative with respect to its maturity. Within all local volatility models,
the Constant Variance of Elasticity (CEV) model [4] attracts most interest:

\[ dS_t = S_t(\mu dt + S_t^{\alpha - 1}dW_t) \]  

(1.4)

2. Another widely used approach is to assume that the volatility process in (1.2) follows a stochastic diffusion process. And thus these kind of models are called stochastic volatility models [5–12]. The stochastic volatility model can be written is the form of:

\[ dS_t = S_t(\mu dt + \sigma_t dW_t) \]  

(1.5)

\[ d\sigma_t = f(\sigma_t, t)dt + g(\sigma_t, t)dZ_t \]  

(1.6)

\[ dW_t dZ_t = \rho dt. \]  

(1.7)

A remarkable result is obtained by Heston [12], in which the following model was proposed:

\[ dS_t = S_t(\mu dt + \sqrt{V_t}dW_t) \]  

(1.8)

\[ dV_t = \kappa(\theta - V_t)dt + \sigma_V \sqrt{V_t}dZ_t \]  

(1.9)

\[ dW_t dZ_t = \rho dt. \]  

(1.10)

The author got the closed form solution to the European call option price, by the Fourier Transform method. This opens the door to more general so-called affine models. For these models, the dynamics of underlying are carefully chosen such that the models encompass certain features of the market, and also the solutions could be obtained in closed form by solving in closed form a system of ordinary differential equations (ODEs), followed by numerical integration. Fast Fourier transform method was also introduced in finance literature [16,18,60–64,73].

Another remarkable model is the SABR model [13], where the authors proposed the model:

\[ dF_t = \sigma_t S_t^\beta dW_t \]  

(1.11)

\[ d\sigma_t = \alpha \sigma_t dZ_t \]  

(1.12)

\[ dW_t dZ_t = \rho dt. \]  

(1.13)

and obtained closed form approximations to European call option price, by the singular perturbation techniques, assuming the ‘volvol’ parameter \( \alpha \) is small. We will focus on this
model later and show our approximation results.

3. Jump-Diffusion/Pure Jump models:

Starting with Merton’s seminal paper [14] and up to the present date, various aspects of
models with jumps have been studied in the academic finance community. In the last decade,
also the research departments of major banks started to accept jump-diffusions and Levy
processes as a valuable tool in their day-to-day modeling. This increasing interest to jump
models in finance is mainly due to the following reasons. First, in a model with continuous
paths like a diffusion model, the price process behaves locally like a Brownian motion and the
probability that the stock moves by a large amount over a short period of time is very small,
unless one fixes an unrealistically high value of volatility. Therefore, in such models the prices
of short term out of the money (OTM) options should be much lower than what one observes
in real markets. On the other hand, if stock prices are allowed to jump, even when time to
maturity is very short, there is a non-negligible probability that after a sudden change in
the stock price the option will move in the money. Second, from the point of view of hedging,
continuous models of stock price behavior generally lead to a complete market or, at least, to
a market, that can be made complete by adding one or two additional instruments, like in
stochastic volatility models [12]. Since in such a market every terminal payoff can be exactly
replicated, options are redundant assets, and the very existence of traded options becomes
a puzzle. The mystery is easily solved by allowing for discontinuities: in real markets, due
to the presence of jumps in the prices, perfect hedging is impossible and options enable the
market participants to hedge risks that cannot be hedged using the underlying only. From
a risk management perspective, jumps allow to quantify and take into account the risk of
strong stock price movements over short time intervals, which appears non-existent in the
diffusion framework. The last and probably the strongest argument for using discontinuous
models is simply the presence of jumps in observed prices. Among others several remarkable
papers are [15,16,18,73].

1.2 Analytical/Numerical Approaches to Heat Kernel/Op-
tion Pricing/Implied Volatility

By the Feynmann Kac formula, the Black Scholes European call option satisfies the PDE:

\[ \partial_t u = \frac{\sigma^2}{2} \partial^2_S u + rS \partial_S u - ru \]  

(1.14)
where \( t \) is the time to maturity. with initial condition

\[
u(S,0) = (S - K)_+.
\]

This partial differential equation (PDE) admits a closed form solution:

\[
u(S,t) = SN(d_1) - Ke^{rt}N(d_2)
\]

where
\[
d_1 = \frac{\log S - \log K + (r + \frac{\sigma^2}{2})t}{\sigma \sqrt{t}},
\]
and
\[
d_2 = d_1 - \sigma \sqrt{t}.
\]

\( N \) is the cumulative distribution function for the normal distribution

\[
N(x) = \int_{-\infty}^{x} \frac{e^{-y^2/2}}{\sqrt{2\pi}}dy.
\]

While there are seldom closed-form solution available for most of the models developed in literature [1, 4, 12, 19], analytical/numerical approximation methods were developed, during the past thirty years. Numerical methods in the PDE level such as finite difference/finite element methods, or in the SDE level such as Monte Carlo simulation methods, are not in the scope of our discussion. We will focus on the closed-form approximation for the solutions. The advantages of getting closed-form approximation are three-folded: first of all, calibration with closed form solution is fast, since one can precompute all the combinations of parameters in advance. This feature is extremely valuable for real time model calibration, where computation speed is critical. Second, the closed form approximation gives us insight of how the parameters effect the option price/implied volatility in a direct way, instead of relying on the black-box of other numerical methods. Last, numerical experiment shows that even for the truncation of few terms of the approximation, one gets satisfactory numerical result. For example, taking into consideration of bid-ask spread and other market microstructures, a price with absolute error within .5 cents, or relative error within \(10^{-5}\) could be regarded as a good approximation.

In the following: we will follow three aspects of closed form approximation of the PDE for the option pricing problem:
1. Approximation to the transition density function
2. Approximation to the option price
3. Approximation to the implied volatility
4. Approximation to the characteristic function

Although these three aspects are closed related, they have their own interest:

The transition probability density (heat kernel, Green’s function, fundamental solution) are of fundamental importance, in the sense that the solution of the linear parabolic equations can be obtained by the convolution of the transition probability density with the initial condition. On the other hand, although for European option, the transition probability density function will not appear directly in the pricing formula, it is inevitable in pricing American style options. Lastly, there are amounts of literatures on approximating the heat kernel [25, 26], which could be/has been used in approximating the option prices/implied volatilities in finance literature [21–24]

Option pricing approximation attracts most interest from researcher/practitioner from the aspect of inference market information from the observed market price, since price is the main observable variable on the market.

The concept of implied volatility is important for trading and is extensively used by traders. Actually, it is common practice on trading floors to quote and to observe prices by the implied volatility. A great advantage of having prices expressed in such dimensionless unit is to provide easy comparison between products with different characteristics [27]. Some quantitative researcher also calibrate the implied volatility surface in practice, rather than calibrating the option prices.

Characteristic function is the Fourier transform of the transition kernel, and it carries all the statistical information of the underlying stochastic processes, from which some important statistical quantities could be produced such as skewness, kurtosis. In fact, the statistical analysis of market data implied skewness and kurtosis which do not agree with Black-Scholes assumption, and motivates researchers to seek other models.
1.2.1 Approximation to the transition probability density

Consider the following second order parabolic equation:

\[
\begin{cases}
\partial_t u(t, x) - Lu(t, x) = g(t, x) & \text{in } (0, \infty) \times \mathbb{R}^N \\
u(0, x) = f(x) & \text{on } \{0\} \times \mathbb{R}^N,
\end{cases}
\] (1.16)

where

\[
Lu(x) := \sum_{i,j=1}^{N} a_{ij} \partial_i \partial_j u(x) + \sum_{k=1}^{N} b_k(x) \partial_k u(x) + c(x) u(x),
\] (1.17)

where \(x = (x_1, \ldots, x_N) \in \mathbb{R}^N\), \(\partial_k := \frac{\partial}{\partial x_k}\), and the coefficients \(a_{ij}, b_i\) and \(c\) and all their derivatives are assumed to be smooth and uniformly bounded.

Under very broad conditions, it can be proved that there is a function \(G(t; x, y)\), such that it solves (1.16), with \(G(0; x, y) = \delta_y(x)\).

Given that \(G(t; x, y)\) arises in several different contexts, we will call the function \(G(t; x, y)\) the transition probability density, pricing kernel, heat kernel, fundamental solution, or Green function interchangeably.

There is a large literature on developing methods to obtain asymptotic formulas for the Green’s function for \(t\) small and \(y\) close to \(x\). For example, interpreting the operator \(L\) as a Laplace-Beltrami operator on a manifold plus lower order terms, leads to the asymptotic expansions of the form

\[
G(t; x, y) = \sum_{j \geq 0} t^{(j-n/2)} p_j(x, t^{-1/2}(x-y)) e^{-(x-y)^T A^{-1}(x)(x-y)/4t} (1.18)
\]
as \(t \to 0_+\), where \(d(x, y)\) is the geodesic distance between \(x\) and \(y\) and \(G^{(j)}(t; x, y)\) are smooth functions in \(x\) and \(y\). Among the vast literature we refer to [25, 26, 28–32]. The results of these works were used in the option pricing literature such as [21–24].

However, one difficulty in the practical implementation of this geometric approach is that in general there are no formulas for the geodesic distance, which thus needs to be accurately approximated or computed numerically.

A related short-time asymptotic approach uses oscillatory type integers, which gives:

\[
G^L(t, x, y) \sim \sum_{j \geq 0} t^{(j-n/2)} p_j(x, t^{-1/2}(x-y)) e^{-(x-y)^T A^{-1}(x)(x-y)/4t} (1.19)
\]
as \(t \to 0_+\), where \(p_j(x, w)\) is a polynomial of degree \(j\) in \(w\), and \(A(x) := [a_{ij}(x)]\). We refer
to [34] Chapter 7, Section 13, where a heat parametric was constructed on compact manifolds. The advantage of this approach is that one does not need to compute the geodesics, which is in general not an easy task for dimensions greater than one. In his book [34], the author proposed a system of transport equations, which in theory could be computed explicitly. However, from the computational point of view, it is still not clear how to implement the algorithm such that one can easily compute arbitrarily higher order terms of \( p_j(x, z) \), either by hand or by computer programming.

Recently, a new method call Dyson-Taylor commutator method was developed in [36], and was applied to option pricing in [37]. Essentially, the authors finds a computable way for constructing the parametric for the elliptic operator \( L \) in (1.16). In these papers, the authors freeze the the variable coefficients at some point \( z \), which is closed related to \( x \) and \( y \), where \( \mathcal{G}(t; x, y) \) is concerned, and then use a Dyson-series perturbative expansion to approximate the heat-kernel of \( L \). The key point is that the Dyson-series expansion turns out to be explicitly computable using the Baker-Campbell-Hausdorff commutator formula. We also refer to several closely related papers [38–42].

In this dissertation, we further develop the Dyson-Taylor commutator method, that we develop a perturbative approach, such that we can compute the higher orders recursively, which on one side gives analytical expression for the higher order terms approximation, and more importantly, enables us to compute the higher order terms in a convenient and effective way by computer programming. We hope this will eventually make the Dyson-Taylor method (or other parametric construction approach) applied to financial industry. The approach we developed in this paper is different from the previous approaches, in that we obtained a new forms of approximations which enabled me to overcome the computational difficulty for arbitrarily high order terms, and by symbolic computation, we got very clean closed form approximations to option price/heat kernel/implied volatility for various models, as shown in the following chapters.

Furthermore, we did not approximate the heat kernel in the form of (1.19), instead, our approximation to the Green’s function is in the form of:

\[
\mathcal{G}^L(t, x, y) \sim e^{\sum_{j \geq 1} \frac{((d-n)/2)}{2} p_j(x, t^{-1/2}(x-y))} e^{-(x-y)^T A^{-1}(x-y)} / \sqrt{4t}
\]

We immediately notice that this keeps the property that the Green’s function is non-negative. We will see later that this expression gives us cleaner results also.

We will use the approximated heat kernel to create American style option pricers later.
We also note that there is a parallel approach in the stochastic differential equation perspective, where the authors either on Malliavin calculus to get expansive approximations, or using the PDE approach, to approximate the transition probability densities, with application to statistical inferences, please refer to [43–47] and references therein.

The following is the result for approximating the heat kernel of the CEV model 1.4. Note that when we are in forward measure, the interest rate \( r = 0 \). This is the general treatment in literature [23].

**Theorem 1.2.1.** The heat kernel for the CEV model is approximated by:

\[
G_{CEV}(t; F_0, F_t) = \exp(\sigma_z^2 t A) \frac{\exp\left(-\frac{(\log F_0 - \log F_t - 5\sigma_z^2 t)^2}{2\sigma_z^2 t}\right)}{\sqrt{2\pi t \sigma_z F_t}}
\]  

where \( \sigma_z = \frac{a}{f_{av}^{1-\beta}} \) and \( f_{av} = \sqrt{F_0 K} \). \( A \) has the following expansive expression:

\[
A = a_1(\beta - 1) + a_2(\beta - 1)^2 + a_3(\beta - 1)^3 + a_4(\beta - 1)^4 + \cdots
\]  

with

\[
\begin{align*}
a_1 &= \frac{k}{T} \\
a_2 &= \frac{T^2}{384} - \frac{T}{16} - \frac{k^4}{24T^2} + \frac{1}{8} \\
a_3 &= 0 \\
a_4 &= -\frac{k^6}{720} - \frac{T(3T^3 - 120T^2 + 1040T - 960)}{15360}
\end{align*}
\]

where \( T = \frac{a^2 t}{f_{av}^{2-2\beta}} \) is the rescaled time to maturity, \( k = \log F_0 - \log K \) is the log-moneyness.

We also have the following results for the transition kernel of SABR model, which can be compared to the results in [24]

**Theorem 1.2.2.** The marginal pdf for the SABR model is approximated by:

\[
G_{SABR}(t; F_0, \sigma_0, F_t, \cdot) = \exp(\sigma_z^2 t A) \frac{\exp\left(-\frac{(\log F_0 - \log F_t - 5\sigma_z^2 t)^2}{2\sigma_z^2 t}\right)}{\sqrt{2\pi t \sigma_z F_t}}
\]  

(1.23)
where $\sigma_z = \frac{\sigma_0}{f_{av}^{1-\beta}}$ and $f_{av} = \sqrt{F_0K}$. $A$ has the following expansive expression:

$$A = a_1(\beta - 1) + a_2(\beta - 1)^2 + a_3(\beta - 1)^3 + a_4(\beta - 1)^4 + \cdots$$  \hspace{1cm} (1.24)

with

$$a_1 = \frac{sk}{T} + \frac{\alpha \rho}{\sigma_z} \left( \frac{T^2}{384} - \frac{T}{16} - \frac{k^4}{24T^2} + \frac{1}{8} \right) + \frac{\alpha^2 \rho^2}{\sigma_z^2} \left( \frac{1}{4} + \frac{T}{48} - \frac{k^2}{12} - \frac{3k^2}{4T} + \frac{k^4}{6T^2} + \frac{T^2}{96} \right)$$

$$a_2 = \frac{sk}{T} + \frac{\alpha \rho}{\sigma_z} \left( \frac{T^2}{384} - \frac{T}{16} - \frac{k^4}{24T^2} + \frac{1}{8} \right) + \frac{\alpha^2 \rho^2}{\sigma_z^2} \left( \frac{1}{4} + \frac{T}{48} - \frac{k^2}{12} - \frac{3k^2}{4T} + \frac{k^4}{6T^2} + \frac{T^2}{96} \right)$$

where $s = \beta - 1$, $T = \frac{a^2 t}{f_{av}^{2-2\beta}}$ is the rescaled time to maturity, $k = \log F_0 - \log K$ is the log-moneyness. $a_3$ and $a_4$ are computable.

### 1.2.2 Approximation to the option price

As one can approximate the fundamental solution, it is straightforward to approximate the solution of the linear parabolic PDE. One might think numerical integration is inevitable, but actually there are no numerical integrations in the closed form approximation. As far as I know, this idea is originally developed in [48], in which the authors approximate the option price for a broad range of models, by the Black-Scholes, plus perturbation terms, which result from taking derivatives of the Black-Scholes formula with respect to the log stock price. Further development are included in [38–42]. Again, approximation for higher order terms are difficult for their approach, since the number of terms increase exponentially with respect to the approximation order.

We do not explicitly write down the approximated option price, since I already got the closed-form approximation of the implied volatility in the next subsection. When substitute to Black-Scholes formula, we recover the approximated option price.
1.2.3 Approximation to the implied volatility

There are vast amount of literature devoting to this topic. As we stated before, basically it is because the traders quote the implied volatility, instead of the option price, and that some of the model calibrations procedures are calibrating the implied volatility surface, instead of calibrating the price surfaces.

There seems several approaches to the approximation to the implied volatility. One straightforward way is that when approximation of the option price is available, one can invert the Black-Scholes formula to recover the expansive approximation to the implied volatility [38–42]. The problem with this approach is that it is computational very complicated, and the reasons are two-folded: first, the approximated option price is expansive expression, which involves the Black-Scholes price and its higher (mathematical) derivatives with respect to the log price. Second, inverting the Black-Scholes to recover the implied volatility is complicated, especially when one wants to get higher order approximations, because one needs to take derivative with respect to volatility in Black-Scholes. The two-folded computational complication makes the resulting expansive expression more complicated, and it is computational extremely intensive to get higher order approximations.

Another approach is taken by Hagan and his coauthors. In their 1999 paper [23], they used the singular perturbation expansion to get closed form approximation to the European call option price for the CEV model, where they expand the option price by the Black-Scholes plus perturbative terms, which are the derivatives of Black-Scholes with respect to time to maturity, and thus making the task of inverting the Black-Scholes to implied volatility easier. Their result has been set to the benchmark for academia/industry until now, and I will present my new result, which is cleaner in expression, more systematic in theory, easy to compute in any higher order with computational codes, and higher precision of accuracy illustrated by numerical experiments. In their 2005 paper [24], Hagan and his coauthors give a expansive approximation for the implied volatility of SABR model. They connected the implied volatility and the local volatility, which is a weighted average of instant volatilities (for more details please refer to [49], where the relationship between local volatility and stochastic volatility models were discussed), and they focused on the approximation of the heat kernel of the SABR model. This approach also gives back the benchmark for literature and industry the implied volatility of SABR model, and we will show later how we get our own way of approximation to the implied volatility and will compare with their results.

A similar approach is taken in [21], where the author applied the results of heat kernel
expansion on Riemannian manifolds. It should be noticed that both Hagan and his coauthors [23, 24], and Henry-Labordere [21] did not construct the parametrix. The advantage is that they gain the geometric flavor in the approximated heat kernel, and it happens to be that the geodesics are easy to compute in closed form; however, they both have difficulties to get higher order approximations, regardless of the necessity of doing so. In fact, even for lower order approximations, it is quite demanding when they solve by hand calculation. On the other hand, parametric construction approach loses the geometric flavor, but gains the computational feasibility, in the sense that we can get arbitrarily higher order terms in the approximation. Both approach have advantages/disadvantages, we will show that for CEV model, our numerical performance is arguably better; while for SABR, we should treat the results case by case.

It is worth to mention another approach to the implied volatility [27, 50, 51], instead of inverting the Black-Scholes, the authors directly considered the PDE the volatility satisfies. However, it is not clear that the resulting nonlinear parabolic equation could be solved efficiently, to get expansive approximation of the implied volatility. The other problem might be that the resulting nonlinear parabolic equation is not as clean as the authors showed in formula (15) in [51], since the computation as shown in formula (26) in the paper might possibly be incorrect.

Our main result for the closed form approximation of implied volatility for CEV model is the following:

**Theorem 1.2.3.**

\[ dF(t) = aF^\beta dW(t) \]  

(1.25)

Then the implied volatility has the expansion:

\[ \sigma^2_B = \frac{a^2}{f_{av}^{2-2\beta}} \left( 1 + (\beta - 1)^2 Q_1 + (\beta - 1)^4 Q_2 + \cdots \right) \]  

(1.26)

or

\[ \sigma_B = \frac{a}{f_{av}^{1-\beta}} \left( 1 + (\beta - 1)^2 Q_1/2 + (\beta - 1)^4 (Q_2/2 - Q_1^2/8) + \cdots \right) \]  

(1.27)

where \( f_{av} = \sqrt{F_0 K} \), \( k = \log(F_0/K) \) is the log-moneyness, and \( T = \frac{a^2 t}{f_{av}^{2-2\beta}} \) is the scaled time to maturity.

\[ Q_1 = \left( \frac{T - k^2}{12} - \frac{T^2}{48} \right) \]
\[ Q_2 = \frac{T^4}{640} + \frac{121T^3}{5760} + \frac{13T^2k^2}{2880} + \frac{43T^2}{1440} + \frac{7Tk^2}{480} + \frac{k^4}{240} \]

The following is a comparison between our approximation and Hagan’s approximation. The benchmark is the closed form solution of CEV model in [20], which involves non-central chi-square distribution, such that it is slow in practice, but accurate enough to give us a reliable values to compare with. The model parameters are:

CEV model with \( \sigma = 30\%, \beta = .75, t = 2, F_0 = 100, r = 0 \), comparison between Hagan’s approximation and fifth order Dyson-Taylor approximation.

<table>
<thead>
<tr>
<th>Strike</th>
<th>75</th>
<th>83.33</th>
<th>91.67</th>
<th>100</th>
<th>108.33</th>
<th>116.67</th>
<th>125</th>
<th>133.33</th>
<th>141.67</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accurate Sol.</td>
<td>25</td>
<td>16.74</td>
<td>9.11</td>
<td>3.57</td>
<td>0.92</td>
<td>0.15</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Err. Hagan(( \times 1E-6 ))</td>
<td>0.21</td>
<td>0.05</td>
<td>0.0005</td>
<td>0.001</td>
<td>0.0008</td>
<td>0.024</td>
<td>0.1098</td>
<td>0.1871</td>
<td>0.1779</td>
</tr>
<tr>
<td>Err. DS(( \times 1E-10 ))</td>
<td>1.0025</td>
<td>0.2184</td>
<td>0.6088</td>
<td>0.5953</td>
<td>0.2613</td>
<td>0.076</td>
<td>0.0053</td>
<td>0.016</td>
<td>0.0154</td>
</tr>
</tbody>
</table>

**Theorem 1.2.4.** The implied variance of SABR model has the expansion:

\[ V_{imp} = \tilde{V}(1 + q_1 + q_2 + \cdots), \]

with

\[
\tilde{V} = V_z
\]

\[
q_1 = \frac{\rho \alpha (T - 2k)}{\sqrt{V_z}}
\]

\[
q_2 = (\beta - 1)^2 \left( \frac{T - k^2}{12} - \frac{T^2}{48} \right) + \frac{\alpha^2 (-T^2 + 2T + 4k^2)}{12V_z} + \frac{\alpha^2 \rho^2 (-4T^2 + 5T^2 - 4k^2 - 8Tk)}{16V_z} - \frac{s\alpha \rho T (T - 12 - 2k)}{24\sqrt{V_z}}
\]

where \( V_z = \frac{\sigma_0^2}{f^2_{av-2\beta}}, f_{av} = \sqrt{F_0K}, k = \log(F_0/K) \) is the log-moneyness, and \( T = \frac{a^2t}{f^2_{av-2\beta}} \) is the scaled time to maturity.
1.2.4 Approximation to the characteristic functions

For a random variable $X$, the characteristic function (CF) is:

$$\Phi_X(u) = \mathbb{E}(\exp(iuX))$$ (1.28)

The cumulant generating function is defined as:

$$\phi_X(u) = \log(\Phi_X(u))$$ (1.29)

The cumulants or semi-invariants are defined as:

$$c_n(X) = \frac{1}{i^n} \frac{\partial^n \phi_X}{\partial u^n}(0)$$ (1.30)

Of all the statistical features of the process, people are most interested in the following:

$$s(X) = \frac{c_3(X)}{c_2(X)^{3/2}}$$ (1.31)

$$\kappa(X) = \frac{c_4(X)}{c_2(X)^2}$$ (1.32)

In empirical studies, people always found that the log price is skewed and have positive kurtosis under risk neutral measure, which contradicts the log-normal assumption of Black-Scholes. New models were developed to captures there feature. However, it seems that there were no quantitative relations between model parameters and skew/kurtosis up to now.

We have the following results:

**Theorem 1.2.5.** For the CEV process

$$dF_t = \sigma F^\beta dW_t$$

The skewness

$$s_{CEV} = 3\sqrt{T}s - \frac{13T^{3/2}}{4}s^2 + \frac{3\sqrt{T}(39T^2 + 80T)}{32}s^3 - \frac{\sqrt{T}(501T^3 + 4464T^4)}{128}s^4 + \ldots$$
The kurtosis

\[ \kappa_{CEV} = 20T s^2 - 45T^2 s^3 + 2T (38T^2 + 67T) s^4 + \cdots \]

**Theorem 1.2.6.**

For the SABR process

\[
\begin{aligned}
    dS_t &= \sigma_t F_t dW_t \\
    d\sigma_t &= \alpha \sigma_t dZ_t
\end{aligned}
\]

with \( dW_t dZ_t = \rho dt \). The skewness

\[ s_{SABR} = 3\sqrt{T}s + 3\sqrt{t}\alpha \rho - 2\sqrt{T}t\alpha^2 - \frac{13T^{3/2}}{4}s^2 + \cdots \]

The kurtosis

\[ \kappa_{SABR} = 20T s^2 + 4t\alpha^2 + 36\sqrt{T}ts\alpha \rho - 45T^2 s^3 + 39Tts\alpha^2 + \cdots \]

where \( s = \beta - 1, T = \sigma^2 F_0^{2(\beta-1)} t \)

Note that for SABR and other stochastic volatility models, the skewness and kurtosis are marginal on the stock price process.

### 1.2.5 When coefficients are time-dependent

The problem when the coefficients also depend on time is of particular interest in practice, since one adds more term structure of the model, such that more satisfactory pricing, hedging and calibration result might be achieved. Within vast literature, we refer to [21,52–58]. In this dissertation, I extended the Dyson-Taylor commutator method such that it could be applied to time-dependent coefficients. This approach is different from [37] by the original authors of Dyson-Taylor method. I used the result from time-dependent coefficient case, to overcome the stiffness brought by the elastic parameter \( \kappa \) in Heston model 1.8, to get the following result:

**Theorem 1.2.7.** For Heston model 1.8, the implied variance of the European call option has the following expansive expression:

\[ V_{imp} = q_0 + q_1 + \cdots \] (1.33)
where $m_0 = s_z + w_0$ and $k = \log(S_0) - \log(K) + rt$ is the log-moneyness, and $V_z$ is the averaged volatility of the Black-Scholes formula with deterministic time-dependent volatilities:

$$V_z = \frac{1}{t} \int_0^t (\theta + \exp(-\kappa \tau)(m_0 - \theta))d\tau$$

$$f(t) = \exp(-t) - 1 + t, \text{ and } g(t) = \exp(t) - 1 - t, \text{ and } w_0 \text{ could be chosen as 0.}$$

### 1.3 Organization of the dissertation

In this dissertation, we further develop the Dyson Taylor method, which was originally developed in order to get a series expansion for the heat kernel in a computational feasible way. By taking one step further, we integrate the approximate heat kernel with the initial condition, to get a series expansion approximation to the solution, for given initial conditions. The approximated implied volatility will also be considered, where we use two different ways to compute the same option price, thus getting an approximated expression for implied volatility for various models. After we treat in general case and established the suitable theoretical framework, we apply the approximation to several examples; specifically, we will show how to approximate the European style option price, under local/stochastic volatility models, including famous Constant Elasticity of Variance (CEV) model, Heston model and SABR model. We also outline the framework to approximate American style options, which corresponds to free boundary value problems. We provide some examples for computing the series expansion approximation for options with jumps, which corresponds to certain type of partial-integral differential equations. Time-dependent coefficients will also be considered in this dissertation.

The dissertation is organized in the following way:

In chapter 2, we get the perturbative expansion with small parameters for second order parabolic equation, with space dependent coefficients. The main tool we will use is the
Baker-Campbell-Hausdorff commutator formula. As a result, most of the following chapters are natural extensions of this first chapter.

In chapter 3, we continue the work of Dyson Taylor commutator method. Specifically, we seek the short-time expansion of heat kernel. This is achieved by ‘parabolic rescaling’, which enables us to treat the time variable as a small parameter, and thus we can use the result of first chapter. The approach is different from the original Dyson Taylor commutator method. Although the results are the same, our approach is more feasible from computational point of view.

In chapter 4, we move one step further, to seek the approximate solution, by integrating the initial condition against the approximated heat kernel we obtained from chapter 3. After realizing that the heat kernel is just a specific solution with delta function as its initial condition, we introduce two different ways to directly approximate the solution of parabolic PDE, with the first one similar but more feasible in computing the implied volatility, and the second method new as far as I know. We will prove that these two new approaches all give us asymptotical solutions under certain regularity conditions.

In chapter 5, we treat higher dimension models, and will show our approximation results for Heston model and SABR model. Again, symbolic computation is carried out, and comparisons are made with results in literature. We also compute the skewness and kurtosis for CEV and SABR models. It is generally considered that SABR model is more flexible to produce volatility skew/smile, and my result here is the first time people can see quantitatively how skew/smile can be produced, with different choice of model parameters.

In chapter 6, we will show how to extend our approximation to time dependent coefficients. Since time dependence gives more flexibility to calibrating models, we think this chapter is important for the popularity of our approximation methods to be used in industry. The basic idea is that we ‘freeze whatever makes trouble’. We will explain this principle further in this chapter, and show our example of calculating the bond price of Hull White model.

In chapter 7, we illustrate the idea to price the American style option, in which we get small time interval for free. We can not expect closed form expansion in such situation, but we can still compute certain models in a fast way. CEV model will again be computed in this chapter, for American put option.

In chapter 8, we will treat the jump-diffusion models. In a very general situation, the parabolic rescaling skill does not apply, because of existence of the integro-differential operators. Nevertheless, we can still use the two other ways we introduced, and the computation of local volatility version of Kou’s/Bates’ model will be shown in this chapter.
In chapter 9, we will focus on the implied volatility approximation of various models. Industry practitioners can directly go to this chapter to find useful approximations formulas for implied volatilities. The results are new, to the understanding of the author.

In chapter 10, we will conclude and give further research directions.

The appendix include the illustration of the MATLAB code, which was used for the computations all through this dissertation.
Chapter 2  |  Formal Expansion for Small Parameter Perturbation

In this section, we derive the formal expansion for a second order parabolic equation, when perturbed by a small parameter. The following sections are natural expansion of this section.

Explicitly, we find the formal solution in regular perturbation form of the following PDE:

$$u_t = L^s u$$  \hspace{1cm} (2.1)

with initial condition:

$$u(0) = f$$  \hspace{1cm} (2.2)

where $L^s$ is a second order partial differential operator with small parameter $s$. We assume that $L(s)$ has the following expansion form:

$$L^s = \sum_{j=0}^{\infty} s^j L_j$$  \hspace{1cm} (2.3)

where $L_j$ is in the form of

$$a_j(x)\partial_x^2 + b_j(x)\partial_x + c_j$$

Please note that we write the PDE in 1d form for the sake of simplicity. We should keep in mind that we are actually treating the problem in any dimension. We will see later how two and three dimensional problems are considered.

This is closely related to the short time expansion of the Green’s function of second order
parabolic equations. Also, the result is of its own interest.

We seek the solution in the form of:

\[ u(t) = \sum_{j=1}^{\infty} s^j u_j(t) \]  \hspace{1cm} (2.4)

By matching the coefficients of \( s \), we find the following relationship:

\[ \frac{\partial u_0}{\partial t} = L_0 u_0 \]  \hspace{1cm} (2.5)

with

\[ u_0(0) = f \]  \hspace{1cm} (2.6)

and

\[ \frac{\partial u_j}{\partial t} - L_0 u_j = \sum_{k=1}^{j} L_k u_{j-k} \]  \hspace{1cm} (2.7)

with

\[ u_j(0) = 0 \]  \hspace{1cm} (2.8)

for \( j \geq 1 \). The solution for \( u_0 \) is clear, but we also need to find the solution for general \( u_j \).

### 2.1 The Commutators

Following [36], we have the following definitions:

**Definition 2.1.1.** (Spaces of Differential Operators) For any nonnegative integers \( a, b \) we denote by \( \mathcal{D}(a, b) \) the vector space of all differential operators of polynomial degree at most \( a \) and order at most \( b \). We extend this definition to negative indices by defining \( \mathcal{D}(a, b) = 0 \) if either \( a \) or \( b \) is negative. By polynomial degree of \( A \) we mean the highest power of the polynomials appearing as coefficients in \( A \).

We remark that \( \mathcal{D}(0, b) \) consist of differential operators with constant coefficients.

**Definition 2.1.2.** (Adjoint Representation). For any two differential operators \( A_1 \) in \( \mathcal{D}(a_1, b_1) \) and \( A_2 \) in \( \mathcal{D}(a_2, b_2) \) we define \( ad_{A_1}(A_2) \) by

\[ ad_{A_1}(A_2) := [A_1, A_2] = A_1 A_2 - A_2 A_1, \]
as usual, and for any integer \( j \geq 1 \) we define \( ad_{A_1}^j(A_2) \) recursively by

\[
ad_{A_1}^j(A_2) := ad_{A_1}(ad_{A_1}^{j-1}(A_2))
\]

The following lemmas are also from [36]:

**Lemma 2.1.3.** Suppose \( A_1 \in D(a_1, b_1) \) and \( A_2 \in D(a_2, b_2) \). Then for any integer \( k \geq 1 \), \( ad_{A_1}^k(A_2) \in D(k(a_1 - 1) + a_2, k(b_1 - 1) + b_2) \)

**Lemma 2.1.4.** Let \( m, k \) be fixed integers \( \geq 1 \). Let \( L_0 \in D(0, 2) \) and \( L_m \in D(m, 2) \). Then, \( ad_{L_0}^k(L_m) \in D(m - k, k + 2) \). In particular,

\[
ad_{L_0}^k(L_m) = 0, \quad \text{if} \quad k > m. \tag{2.9}
\]

**Lemma 2.1.5.** *(Baker-Campbell-Hausdorff)* Let \( L_0, L_m \in \mathbb{L}_\gamma \). Then

\[
e^{tL_0}L_m = e^{t ad_{L_0} L_m e^{tL_0}} = \left( \sum_{k=0}^{\infty} t^k ad_{L_0}^k(L_m)/k! \right) e^{tL_0}.
\]

Remark: If \( L_0 \in D(0, 2) \cap \mathbb{L}_\gamma \) and \( L_m \in D(m, 2) \), then the infinity summation in the previous lemma is actually finite by Lemma 2.1.4.

**Lemma 2.1.6.** Let \( L_0 \in D(0, 2) \cap \mathbb{L}_\gamma \), and let \( L_m \in D(m, 2) \). Then for any \( \theta > 0 \),

\[
e^{\theta L_0}L_m = P_m(L_0, L_m; \theta, x, \partial)e^{\theta L_0} \tag{2.10}
\]

where \( P_m(\theta) = P_m(L_0, L_m; \theta, x, \partial) \in D(m, m + 2) \) is given by

\[
P_m(\theta) := \sum_{k=0}^{m} \frac{\theta^k}{k!} ad_{L_0}^k(L_m) = L_m + \theta[L_0, L_m] + \frac{\theta^2}{2}[L_0, [L_0, L_m]] + \cdots. \tag{2.11}
\]

**Proof.** Apply the Baker-Campbell-Hausdorff formula \( \square \)

### 2.2 Application of the Commutators to the Approximation Problem

Let us write \( u_n(t) \) in (2.7) into the form

\[
u_n(t) = \mathcal{P}_n(t)e^{tL_0}f, \tag{2.12}
\]
where $P_n(t)$ is a differential operator with polynomial coefficients. We also define $P_0(t) = Id$. Then we have the following:

**Lemma 2.2.1.**

\[
P_n(t) = \sum_{l=1}^{n} \int_0^t e^{(t-\tau)\text{ad}_{L_0}(L_lP_{n-l}(\tau))}d\tau \tag{2.13}
\]

**Proof.** From (2.7), we use Duhamel’s principle to get

\[
u_n = \int_0^t e^{(t-\tau)L_0} \sum_{l=1}^{n} L_lu_{n-l}(\tau)d\tau
\]

\[
= \sum_{l=1}^{n} \int_0^t e^{(t-\tau)L_0} L_lu_{n-l}(\tau)d\tau
\]

\[
= \sum_{l=1}^{n} \int_0^t e^{(t-\tau)L_0} L_lP_{n-l}(\tau)e^{\tau L_0} f d\tau
\]

Now we apply Dyson-Taylor commutator to the above to get

\[
u_n = \sum_{l=1}^{n} \int_0^t e^{(t-\tau)\text{ad}_{L_0}(L_lP_{n-l}(\tau))}e^{(t-\tau)L_0} e^{\tau L_0} f d\tau
\]

\[
= \sum_{l=1}^{n} \int_0^t e^{(t-\tau)\text{ad}_{L_0}(L_lP_{n-l}(\tau))} e^{\tau L_0} f d\tau
\]

\[
= \left( \sum_{l=1}^{n} \int_0^t e^{(t-\tau)\text{ad}_{L_0}(L_lP_{n-l}(\tau))} d\tau \right) e^{t L_0} f
\]

\[
\square
\]

For the description of the following theorem, we introduce some notations for the sake of simplicity:

$I_k$ represents all ordered sequences of non-negative integers $(i_1, i_2, \cdots, i_k)$;

$J^*_k$ represents all ordered sequences of positive integers $(j_1, j_2, \cdots, j_k)$;

$m + I_k$ represents all ordered sequences of non-negative integers $(m, i_1, i_2, \cdots, i_k)$ for $m \geq 0$;

$l + J^*_k$ represents all ordered sequences of positive integers $(l, j_1, j_2, \cdots, j_k)$ for $l \geq 1$;

$|I_k| := \sum_{l=1}^{k} i_l; \quad |J^*_k| := \sum_{l=1}^{k} j_l$;

$ad_{L_0}^{i_k}L_{J^*_k} := ad_{L_0}^{i_1}L_{j_1} \cdots ad_{L_0}^{i_k}L_{j_k}$.

Actually, the last definition is ambiguous because the Poisson bracket (commutator) allows no associativity. We define the operation order from left to right, i.e. $ad_{L_0} L_1 ad_{L_0} L_2 =$
\[ \text{ad}_{L_0}(L_1(\text{ad}_{L_0} L_2)). \]

Let’s write the following lemma for later use:

**Lemma 2.2.2.**

\[ \int_0^t (t - \tau)^m \tau^n d\tau = \frac{m! n!}{(m + n + 1)!} t^{m+n+1} \tag{2.14} \]

for any non-negative integers \( m \) and \( n \).

**Proof.**

\[
\begin{align*}
\int_0^t (t - \tau)^m \tau^n d\tau &= t^{m+n+1} \int_0^1 (1 - x)^m x^n dx \\
&= t^{m+n+1} B(m + 1, n + 1) \\
&= t^{m+n+1} \frac{\Gamma(m + 1) \Gamma(n + 1)}{\Gamma(m + n + 2)}
\end{align*}
\]

Here we used the relation between the Beta function and the Gamma function, i.e.

\[ B(m, n) = \frac{\Gamma(m) \Gamma(n)}{\Gamma(m + n)}, \]

and \( \Gamma(m + 1) = m! \) for any non-negative integer \( m \).

Now we are ready to our main result:

**Theorem 2.2.3.**

\[ P_n(t) = \sum_{k=1}^\infty \sum_{|I_k|=-n}^{\infty} \sum_{|I_k|=0}^{\infty} \text{ad}_{L_0} L_{J^*_k} t^{|I_k|+k} \]

\[ \frac{\text{ad}_{L_0} L_{J^*_k}}{(|I_k| + k)!} \]

\[ \tag{2.15} \]

**Proof.** One can check that it’s true for \( n = 0, 1 \). Let’s assume that (2.15) is true for \( n \leq N \), then using Lemma 2.2.1, we have

\[
\begin{align*}
P_{N+1}(t) &= \sum_{l=1}^{N+1} \int_0^t e^{(t-\tau)\text{ad}_{L_0}}(L_l P_{N+1-l} (\tau)) d\tau \\
&= \sum_{l=1}^{N} \int_0^t e^{(t-\tau)\text{ad}_{L_0}}(L_l P_{N+1-l} (\tau)) d\tau + \int_0^t e^{(t-\tau)\text{ad}_{L_0}}(L_{N+1} \text{Id}) d\tau \\
&:= A + B
\end{align*}
\]
Now we use (2.15) for \( n \leq N \), and we get:

\[
A = \sum_{l=1}^{N} \int_{0}^{t} e^{(t-\tau)ad_{L_0}} \left( L_{l} \sum_{k=1}^{N+1-l} \sum_{|J^*_k|=N+1-l} \sum_{|I_k|=0}^{\infty} \frac{ad_{L_0}^{l} J^*_k L_{J^*_k}}{(|I_k| + k)!} d\tau \right)
\]

\[
= \sum_{l=1}^{N} \sum_{k=1}^{N+1-l} \int_{0}^{t} \int_{0}^{\infty} \frac{m!}{(m+|I_k| + k+1)!} \frac{(t-\tau)^m}{m!} d\tau \]

\[
= \sum_{k=1}^{N} \sum_{l=1}^{N+1-k} \sum_{m=0}^{\infty} \frac{(m+|I_k| + k+1)!}{m!} \frac{t^{m+|I_k| + k+1}}{(m+|I_k| + k+1)!} \]

We have used Lemma 2.2.2 and changed the order of summation for the last equality. Let \( K = k + 1 \) in the last equation, if we view \( l \) as a new term added into \( J^*_k \), \( m \) as a new term added into \( I_k \). i.e. \( J^*_K = (l,j_1, \cdots, j_k) \), \( I_K = (m, i_1, \cdots, i_k) \), we get

\[
A = \sum_{K=2}^{N+1} \sum_{|J^*_k|=N+1} \sum_{|I_k|=0} \frac{ad_{L_0}^{l} J^*_k L_{J^*_k}}{(|I_k| + K)!} t^{|I_k| + K}
\]  \hspace{1cm} (2.16)

Now let us look at \( B \)

\[
B = \int_{0}^{t} e^{(t-\tau)ad_{L_0}} (L_{N+1}Id) d\tau
\]

\[
= \sum_{m=0}^{\infty} \int_{0}^{t} \frac{(t-\tau)^m}{m!} ad_{L_0}^{m} L_{N+1}Id d\tau
\]

\[
= \sum_{m=0}^{\infty} \frac{ad_{L_0}^{m} L_{N+1}}{(m+1)!} t^{m+1}
\]

We can write \( B \) as

\[
B = \sum_{K=1}^{1} \sum_{|J^*_k|=N+1} \sum_{|I_k|=0} \frac{ad_{L_0}^{l} J^*_k L_{J^*_k}}{(|I_k| + K)!} t^{|I_k| + K}
\]  \hspace{1cm} (2.17)

Combining (2.16) and (2.17), we have completed the proof.

\]

So we get the following theorem:
Theorem 2.2.1. The formal expansion solution to the problem:

\[ \partial_t u = \sum_{j=0}^{\infty} s^j L_j u \]

is given by:

\[ u(t, x) = \sum_{j=0}^{\infty} s^j \mathcal{P}_n(t) u^0(t, x) \]

where \( u^0(t, x) \) solves:

\[ \partial_t u^0(t, x) = L_0 u^0(t, x) \]

and \( \mathcal{P}_j \) are given inductively by Lemma 2.2.1, or explicitly by Theorem 2.2.3.

2.3 An Example: Small Parameter Expansion for the CEV Model

The Constant Elasticity of Variance (CEV) model utilizes rather simple dynamics for the forward price but the option price has no closed-form solution. Some might argue that in [Schroder-1989], a ‘closed-form’ solution was obtained. However, it relies on the computation of cumulative function of non-central chi-square distribution, and it turns out to be a numerical solution to certain ODE, or an infinitely sum where truncation is needed. In practice, that closed form solution is extremely slow, especially when \( \alpha \) is close 1. This makes it a useful example for the method developed here. For the sake of simplicity, we assume \( r = 0 \); otherwise, all arguments go through with few changes.

Consider the CEV model in the forward measure:

\[ dF_t = \sigma F_t^\alpha dW_t \] (2.18)

where \( \sigma \) is the volatility, \( \alpha \) is between 0 and 2 (although most practitioners are interested in the case \( .5 < \alpha < 1 \); notice when \( \alpha = 1 \), this model is reduced to Black model).

By the Feynman-Kac formula, the option pricing problem is to solve the following PDE:

\[ \partial_t u = .5\sigma^2 F^{2\alpha} \partial^2_u \] (2.19)
with the boundary condition:

\[ u(0, x) = \max(F - K, 0), \]

which is the payoff function for a European Call option. In order to get connected with
the Black-Scholes formula, we first use the logarithm coordinate change for the stock price
variable: Letting \( x = \log F \), and \( v(t, x) = u(t, e^x) \), we have:

\[ \partial_t v = .5\sigma^2 e^{2(\alpha - 1)x} (\partial_x^2 - \partial_x) v \]  \hspace{1cm} (2.20)

Furthermore, letting \( s = 2(\alpha - 1) \), and keeping in mind that \( s \) is the small parameter that we
want to expand for. We will see later that \( s \) does not have to be very small in this numerical
experiment, to give a highly accurate approximation. Given the forward price is \( F_0 \), we are
interested in solving \( v(t, x_0) \), where \( x_0 = \log F_0 \). We rescale the time for the (2.20) by letting

\[ w(t, y) = v\left(\frac{t}{\sigma^2 \exp(s x z)}, y\right) \]  \hspace{1cm} (2.21)

where \( z_0 \) is a ‘frozen’, which we will specify later. We have

\[ v(t, y) = w(\sigma^2 \exp(s x z)t, y) \]

Then \( w \) satisfies the following:

\[ \partial_t w = .5e^{s(x - x_z)}(\partial_x^2 - \partial_x)w \]  \hspace{1cm} (2.22)

Now we do the formal expansion for the equation above, to get:

\[ \partial_t w = .5 \sum_{j=0}^{\infty} s^j \frac{(x - x_z)^j}{j!} (\partial_x^2 - \partial_x) w \]  \hspace{1cm} (2.23)

\[ L_j = \frac{(x - x_z)^j}{2j!} (\partial_x^2 - \partial_x). \]

Next we use the result in the previous section to get:

\[ w(t, x) = \sum_{j=0}^{\infty} s^j P_j(t) w_0(t, x) \]  \hspace{1cm} (2.24)
where $w_0(t, x)$ solves the PDE:

$$\partial_t w_0 = .5(\partial_x^2 - \partial_x)w_0$$

(2.25)

with the boundary condition $w_0(0, y) = \max(e^y - K, 0)$, such that $w_0(t, x) = e^x N(d_1) - KN(d_2)$, where $d_1 = (x - \log K + .5t)/\sqrt t$, and $d_2 = d_1 - \sqrt t$.

The symbolic computation code gives the first few terms of $P_j(t)$; note that we only collect those terms without $(x - x_z)$,

$$P_1(t) = .25(\partial_y^3 - \partial_y^2)t^2$$

$$P_2(t) = .125(\partial_y^2 - \partial_y)t^2 + (5\partial_y^4 - 7\partial_y^3 + 2\partial_y^2)t^3/24 + (\partial_y^6 - 2\partial_y^5 + \partial_y^4)t^4/32$$

$$P_3(t) = (4\partial_y^3 - 5\partial_y^2 + \partial_y)t^3/16 + \ldots$$

We recover the CEV option price by getting:

$$u(t, F_0) = w(\sigma^2 \exp(s x_z) t, \log(F_0)) = w(\sigma^2 \exp(s x_z) t, x_0)$$

(2.26)

And we further substitute the expression of $w(t, x_0)$ in (2.24); after some simplification, we get:

$$Call_{CEV}(F_0, r, t, \sigma, K) = \sum_{j=0}^{\infty} s^j P_j(T) Call_{BS}(\exp(y), r, t, \tilde{\sigma}, K)|_{y=\log F_0}$$

(2.27)

note that $s = 2(\alpha - 1)$, $T = \sigma^2 \exp(s x_z) t$, $\tilde{\sigma} = \sigma \exp(.5 s x_z)$. Numerical tests shows that $x_z = (\log F_0 + \log K)/2$ is the best choice. We will see later why this is true, when we approximate the implied volatilities. Thus, we have completely solve the problem of the small parameter expansion for the CEV model when $\alpha$ is close to 1. Also notice that when we take derivatives of call option w.r.t. the log price, we keep getting closed form expressions, which can be represented by the Hermite polynomials. We skip the details of the computation at this stage, and are willing to give more details when necessary.

We include a numerical experiment for our computation:

From the numerical experiment, we found that the closed form solution to the CEV model is slow for $\alpha$ close to 1. However, our computational time for our method is almost constant for different values of $\alpha$. So we actually provide a method to complement the closed form pricer for CEV, and when $\alpha$ is far from 1, the closed form is accurate as well as fast, for $\alpha$ closed to 1, our method is faster, and almost as accurate as the closed form pricer.
Figure 2.1: CEV model for $\alpha = .85, S_0 = 100, \sigma = .34, t = .75, r = .05$
Chapter 3  
Small Time Expansion for the Heat Kernel-Reformulate the Dyson-Taylor Commutator method

The aim of this chapter is to derive explicitly computable short-time asymptotic expansions for the Green’s function of uniformly parabolic second-order operators with variable, but time-independent, coefficients. This approach is obtained the same approximated series with the original Dyson-Taylor method. However, with different method, we are able to express the series in the inductive matter, which is essential for symbolic computation.

More precisely, we consider the class \( \mathbb{L} \) of second-order differential operators \( L \) with smooth, uniformly bounded coefficients:

\[
Lu(x) := \sum_{i,j=1}^{N} a_{ij} \partial_i \partial_j u(x) + \sum_{k=1}^{N} b_k(x) \partial_k u(x) + c(x) u(x),
\]

(3.1)

where \( x = (x_1, ..., x_N) \in \mathbb{R}^N \), \( \partial_k := \frac{\partial}{\partial x_k} \), and the coefficients \( a_{ij}, b_i \) and \( c \) and all their derivatives are assumed to be smooth and uniformly bounded.

We further assume that \( L \) is uniformly strongly elliptic, namely that there exists a constant \( \gamma > 0 \) such that

\[
\sum_{i,j} a_{ij}(x) \xi_i \xi_j \geq \gamma \|\xi\|^2, \quad \|\xi\|^2 = \sum_{i=1}^{N} \xi_i^2,
\]

(3.2)

for all \( (\xi, x) \in \mathbb{R}^N \times \mathbb{R}^N \). We define the matrix \( A(x) := [a_{ij}(x)] \), which, without loss of generality, to be symmetric. The set of operators \( L \in \mathbb{L} \) satisfying (3.2) will be denoted by \( \mathbb{L}_\gamma \).
We study the short time asymptotic of the initial value problem (IVP) for the parabolic operator $\partial_t - L$,

\[
\begin{aligned}
&
\begin{cases}
\quad \partial_t u(t, x) - Lu(t, x) = g(t, x) & \text{ in } (0, \infty) \times \mathbb{R}^N \\
\quad u(0, x) = f(x) & \text{ on } \{0\} \times \mathbb{R}^N,
\end{cases}
\end{aligned}
\tag{3.3}
\]

for $u, f, g$ in suitable function spaces. In view of Duhamel’s principle, we may assume $g = 0$.

Under certain growth conditions on $u$ and $f$, the unique solution of the IVP (3.3) is obtained by the so-called Green’s function, or fundamental solution or heat kernel $G^L \in C^\infty((0, \infty) \times \mathbb{R}^N \times \mathbb{R}^N)$ such that

\[
u(t, x) = \int_{\mathbb{R}^N} G^L(t, x, y)f(y)dy, \quad t > 0,
\]

We will often write $G^L(t, x, y) = G^L_t(x, y)$. In case we have uniqueness, we shall also use the notation $u(t) = e^{tL}f$. The operator $e^{tL}$ is then called the solution operator of the problem, and its kernel $G^L_t$ the Green’s function, or fundamental solution of $L$, or conditional probability density in applications to probability.

For $L$ with constant coefficients and for a few other cases, one can explicitly compute the kernel $G^L$. In general however, it is not known how to provide explicit formulas for $G^L$, though there is a large literature on developing methods to obtaining good asymptotic formulas for the Green’s function for $t$ small and $x$ close to $y$. A related short-time asymptotic approach uses oscillatory type integrals, which gives:

\[
G^L(t, x, y) \sim \sum_{j \geq 0} t^{(j-n/2)} p_j(x, t^{-1/2}(x - y)) e^{-(x-y)^T A^{-1}(x)(x-y)/4t}
\]

as $t \to 0_+$, where $p_j(x, w)$ is a polynomial of degree $j$ in $w$, and $A(x) := [a_{ij}(x)]$.

In this chapter, we use the idea of the Dyson-Taylor commutator method developed in [36], which based on dilating the coefficients of $L$ around a point $z$ with ration $t^{1/2}$. However, we take a different approach in computing the series expansion, by the regular perturbation of small parameters which was developed in last chapter. The advantage is that we get an explicit expression of the short time expansion for the Green’s function.
3.1 Local Dilations and Perturbative Expansions

Following [36], for any $s > 0$, we consider the action on functions of dilating $x$ by $s$ about $z$ and $t$ by $s^2$ about 0. If $f : \mathbb{R}^N \rightarrow \mathbb{R}$, $u : [0, \infty) \times \mathbb{R}^N \rightarrow \mathbb{R}$, we then set

$$f^{s,z}(x) := f(z + s(x - z))$$
$$u^{s,z}(t, x) := u(s^2t, z + s(x - z))$$

and

$$L^{s,z} := \sum_{i,j=1}^{N} a_{ij}^{s,z}(x)\partial_i \partial_j + s \sum_{i=0}^{N} b^{s,z}(x) + s^2 c^{s,z}(x).$$

We immediately see that

$$L^{s,z}u^{s,z} = s^2 (Lu)^{s,z}, \quad (\partial_t - L^{s,z})u^{s,z} = s^2 [(\partial_t - L)]u^{s,z}$$

In particular, we have the following simple lemma, which we record for further reference

**Lemma 3.1.1.** Let $u$ solve

$$\begin{aligned}
\begin{cases}
\partial_t u(t, x) - Lu(t, x) = 0 & \text{in } (0, \infty) \times \mathbb{R}^N \\
u(0, x) = f(x) & \text{on } \{0\} \times \mathbb{R}^N,
\end{cases}
\end{aligned}$$

then $u^{s,z}$ solves

$$\begin{aligned}
\begin{cases}
\partial_t u^{s,z} - L^{s,z}u^{s,z} = 0 & \text{in } (0, \infty) \times \mathbb{R}^N \\
u^{s,z} = f^{s,z} \in C_0^\infty(\mathbb{R}^N) & \text{on } \{0\} \times \mathbb{R}^N,
\end{cases}
\end{aligned}$$

Now we find the relation between the Green’s function $G_L^t$ and $G_L^{s,t}$. By the definition of the Green’s function, we have

$$u^{s,z}(t, x) = \int_{\mathbb{R}^N} G_L^{s,z}(x, y)f^{s,z}(y)dy$$

$$= \int_{\mathbb{R}^N} G_L^{s,z}(x, y)f(z + s(y - z))dy$$

$$= s^{-N} \int_{\mathbb{R}^N} G_L^{s,z}(x, z + \frac{y - z}{s})f(y)dy$$

On the other hand,

$$u^{s,z}(t, x) = u(s^2t, z + s(x - z)) = \int_{\mathbb{R}^N} G_{s^2t}^{z}(z + s(x - z), y)f(y)dy,$$
which implies
\[
\mathcal{G}_t^L(x, z + \frac{y - z}{s}) = s^N \mathcal{G}_{s^2 t}^L(z + s(x - z), y) \iff \mathcal{G}_t^{L, z}(x, y) = s^N \mathcal{G}_{s^2 t}^L(z + s(x - z), z + s(y - z))
\]

In other words
\[
\mathcal{G}_t^L(x, y) = s^{-N} \mathcal{G}_{s^{-2} t}^{L, z}(z + s^{-1}(x - z), z + s^{-1}(y - z)) \quad \text{for any } s > 0
\]

Letting \( s = \sqrt{t} \), we obtain the desired correspondence between \( \mathcal{G}_t^L \) and \( \mathcal{G}_1^{L, z} \):

**Lemma 3.1.2.** Assume \( L \in \mathcal{L} \), for any \( s > 0 \),
\[
\mathcal{G}_t^L(x, y) = t^{-\frac{N}{2}} \mathcal{G}_1^{L, \sqrt{t}, z}(z + t^{-\frac{1}{2}}(x - z), z + t^{-\frac{1}{2}}(y - z))
\]

From this lemma, we see that since \( \sqrt{t} \) acts as a small parameter for the second order parabolic operator \( L \), one can apply the previous small parameter method to get the approximating formula.

Next we apply Theorem 2.2.3 as well as Lemma 3.1.2 to get a closer look at the approximating Green’s function.

Now let us consider the Taylor expansion of the operator \( L^{s, z} \) around \( s = 0 \),
\[
L^{s, z} = \sum_{m=0}^{\infty} s^m L^z_m \quad (3.4)
\]

The operators \( L^z_m \) are given by
\[
L^z_m := \frac{1}{m!} \left( \frac{d^m}{ds^m} L^{s, z} \right) |_{s=0} \quad (3.5)
\]

The first few terms of the Taylor expansions are explicitly:
\[
L^z_0 = \sum_{i,j=1}^{N} a_{ij}(z) \partial_i \partial_j,
\]
\[
L^z_1 = \sum_{i,j=1}^{N} ((x - z) \cdot \nabla a_{ij}(z)) \partial_i \partial_j + \sum_{i=1}^{N} b_i(z) \partial_i,
\]
\[
L^z_2 = \sum_{i,j=1}^{N} \frac{1}{2} ((x - z)^T \nabla^2 a_{ij}(z)(x - z)) \partial_i \partial_j + \sum_{i=1}^{N} ((x - z) \cdot \nabla b_i(0)) \partial_i + c(z).
\]
where
\[ x \cdot \nabla \phi = \sum_{i=1} x_i \cdot \partial_i \phi \]
and
\[ x^T \nabla^2 \phi x = \sum_{i,j=1} x_i x_j \partial_i \partial_j \phi \]

**Lemma 3.1.3.**
\[ L_k \in \mathcal{D}(k, 2) \] \hspace{2cm} (3.6)

Since \( L_0 \) has coefficients that are constant in \( x \), we obtain
\[ e^{tL_0^z} = \frac{1}{\sqrt{(4\pi t)^N \det A^0}} e^{-\frac{(x-y)^T(A^0)^{-1}(x-y)}{4t}} \]
where
\[ A^0 = [a_{ij}(z)] \] \hspace{2cm} (3.7)

Now one can use the result from last section, to get the following:

**Theorem 3.1.4.** The Green’s function of the operator \( L^s z \) at \( t = 1 \) is:
\[ G_1^{L^s z}(x, y) = \sum_{n=0}^{\infty} s^n \mathcal{P}_n(1)e^{L_0^z} \]
\[ = \sum_{n=0}^{\infty} \sum_{k=1}^{\infty} \sum_{|I_k|=n} \sum_{|J_k^*|=0} s^n \frac{\text{ad}^*_{L_0^z}L^z J_k^*}{(|I_k| + k)!} e^{L_0^z} \]

**Proof.** Compared with Theorem 2.2.3, we only need to prove that for any \( k \in \mathbb{N} \), if \(|I_k| > |J_k^*|\), then \( \text{ad}^*_{L_0^z}L_{J_k^*} = 0 \). This is the result when one use the Lemma 2.1.4 and Lemma 3.1.3 to get the result that \( \text{ad}^*_{L_0^z}L_{J_k^*} \in \mathcal{D}(J_k^* - |I_k|, M) \) for some integer \( M \). By our definition, this term is zero. \( \square \)

Now let us give out the explicit formula for some \( \mathcal{P}_n \). For \( n = 1 \),
\[ \mathcal{P}_1(1) = L_1^z + \frac{1}{2}[L_0^z, L_1^z] \] \hspace{2cm} (3.8)
for \( n = 2 \),

32
\[ \mathcal{P}_2(1) = (L^z + \frac{1}{2} [L_0^z, L^z_0] + \frac{1}{6} [L_0^z, [L_0^z, L^z_0]]) + \frac{1}{2} (L^z_i)^2 \\
+ \frac{L^z_1 [L_0^z, L^z_1]}{6} + \frac{L^z_0, (L^z_0)^2}{24} + \frac{L_0^z [L_0^z, (L^z_0)^2]}{24} + \frac{L_0^z, L^z_1 [L_0^z, L^z_1]}{24} \]

All the above results are the same by the calculation of the iterative method, which is given by [?]. Now let us get something new:

\[ \mathcal{P}_3(1) = L^z_3 + \frac{[L^z_0, L^z_3]}{2} + \frac{[L^z_0, [L^z_0, L^z_3]]}{6} + \frac{[L^z_0, [L^z_0, [L^z_0, L^z_3]]]}{24} \\
+ \frac{L^z_1 L^z_2 + L^z_2 L^z_1}{2} + \frac{L^z_1 [L^z_0, L^z_2]}{6} + \frac{L^z_0, L^z_1 [L^z_0, L^z_2]}{24} + \frac{L^z_2 [L^z_0, L^z_1]}{24} + \frac{L^z_0, L^z_2 L^z_1}{24} \\
+ \frac{[L^z_0, L^z_1 [L^z_0, [L^z_0, L^z_2]]]}{120} + \frac{[L^z_0, [L^z_0, [L^z_0, L^z_2]]]}{120} \\
+ \frac{(L^z_i)^3}{6} + \frac{L^z_1 [L^z_0, (L^z_i)^2]}{24} + \frac{L^z_0, (L^z_i)^2 [L^z_0, L^z_i]}{24} \\
+ \frac{[L^z_0, [L^z_0, (L^z_i)^2]]}{120} + \frac{L^z_1 [L^z_0, (L^z_i)^2]}{120} \\
+ \frac{[L^z_0, L^z_1 [L^z_0, (L^z_i)^2]]}{720} + \frac{[L^z_0, L^z_1 [L^z_0, (L^z_i)^2]]}{720} \]

One can expect that the terms become complicated in exponential order. However, since our formula is explicit, the computer program will definitely help if we want to find the explicit expression for any high order terms.

By substituting this formula to the original Green’s function, one gets:

**Theorem 3.1.5.** The Green’s function of \( L \), when expanding at \( z \), is given by:

\[ \mathcal{G}_i^L(x, y) = t^{-\frac{N}{2}} \mathcal{G}_i^{L^{\sqrt{\tau}_z}}(z + t^{-\frac{1}{2}} (x - z), z + t^{-\frac{1}{2}} (y - z)) \] (3.9)
\[
= \frac{t^{-N/2}}{\sqrt{(4\pi)^N \text{det} A^0}} \sum_{n=0}^{\infty} \sum_{k=1}^{N} \sum_{|I_k|=n} \sum_{|J_k|=0} t^{1/2} a d I_k^L J_k^L e^{-\frac{(x-y)^T(A^0)^{-1}(x-y)}{4t}} \bigg|_{x+\frac{t}{2} (x-z), y+\frac{t}{2} (y-z)}
\]

where \(A^0\) is given by (3.7), and \(L^z_k\) are given by (3.5).

This expression seems complicated, but in my point of view, it is similar to the Taylor expansion for smooth functions.

Actually from this explicit formation, we have the following observations:

We have the following lemma:

**Lemma 3.1.6.** \(P_n(1)\) is composed with terms in \(D(a,b)\) with

\[
a - b \equiv n \pmod{2}
\]

**Proof.** By repeatedly using Lemma 2.1.4 Lemma 3.1.3, one get that every term in \(P_n(1)\) is in the form of \(D(n - |I_k|, 2 + |I_k|)\), where \(|I_k| \leq n\). \(\square\)

**Theorem 3.1.7.** The polynomials \(p_j(z, t^{1/2}(x - z), t^{1/2}(x - y))\) in (3.9) contain only integer powers of \(t\) for every integer \(j\).

**Proof.** Without loss of generality, we assume \(z = 0\) and write \(L_s\) as \(L^{s,z}\). Recall Lemma 3.1.2, we have

\[
G^s(x,y) = t^{-N/2} G^s_1(t^{-1/2} x, t^{-1/2} y)
\]

By Theorem 2.2.3,

\[
G^s_1(x,y) = \sum_{i=0}^{k} s^i P_i(1, x, \partial_x) G^0_i(x,y) + o(s^k),
\]

where \(G^0_i(x, y)\) is the fundamental solution to the unperturbed problem, i.e.,

\[
G^0_i(x,y) = \frac{1}{\sqrt{(4\pi t)^N \text{det} A(0)}} e^{\frac{(x-y)^T A^{-1}(0)(x-y)}{4t}}
\]
Now from Lemma 3.1.6, we see that every single term of
\[ \mathcal{P}_i(1, x, \partial_x) G^0_i(x, y) \]
has the form of
\[ cx^a(x - y)^b G^0_i(x, y) \]
with
\[ a - b \equiv i \pmod{2}. \]
Now we substitute
\[ s = \sqrt{t}, \quad x = t^{-\frac{1}{2}} x', \quad y = t^{-\frac{1}{2}} y' \]
into the term
\[ t^{-\frac{N}{2}} s^i \mathcal{P}_i(1, x, \partial_x) G^0_i(x, y), \]
we get that every term of the above expression is in the form of
\[ cs^{a+b-i} x'^a(x' - y')^b G^0_i(x, y), \]
where
\[ a + b - i \equiv 0 \pmod{2}, \]
which proves the theorem.

By the observation of Lemma 3.1.6, we also get the following lemma:

**Lemma 3.1.8.**
\[ \mathcal{P}_n(1) \in \mathcal{D}(n, 3n) \]

**Proof.** From the proof of Lemma 3.1.6, we have that every term in \( \mathcal{P}_n(1) \) is in the form of
\[ \mathcal{D}(n - |I_k|, 2 + |I_k|), \]
where \( |I_k| \leq n \), which directly gives the lemma. \( \square \)

We can prove the following theorem:

**Theorem 3.1.9.**
\[ \mathfrak{B}^I(z, x, y) = \sum a_{\alpha, \beta}(z)x^\alpha(x - z)^\beta(x - y)^\beta \]
(3.13)
with \( |\alpha| < l, |\beta| < 3l, a_{\alpha, \beta} \in C^\infty_b(\mathbb{R}^N) \)
Proof. From the proof of Theorem 3.1.7, we know that \((x - z)\) terms come from the coefficients of the operators in \(P_l(1)\), while the \((x - y)\) terms are from taking the derivatives of the fundamental solution. By Lemma 3.1.8, we finish the proof this theorem.

Note: this estimation for the order of \(\mathfrak{B}^l(z, x, y)\) is more accurate than Corollary 4.8 of [36]. Further, this result is sharp.

3.2 Convergence of the short time expansion in Certain function space

We explicitly calculated the short-time expansion of Green’s function in last section, and proved some properties of the expansion series. However, we need to establish the convergence result for the formal expansions in our result. We note that same kind of expansion is established in the paper [36], in which the expansion is based on the iterative method by repeatedly using the Duhamel’s principle. In that paper, the convergence result was proved based on the iterative expansion. In this section, we will prove the equivalence between our perturbative expansion and the iterative expansion, such that we can use directly the convergence result in their paper. Even we get the same expansion result, our method is explicit, which can potentially be programmed to give higher approximation in any order.

3.3 Duhamel’s principle and the Iterative Approach

Recall

\[ w_t = Lw \]  

(3.14)

where \(L = \sum_{j=0}^{\infty} s^jL_j\) and \(w(0) = f\).

Theorem 3.3.1. Duhamel’s Principle: The solution to the problem (3.3) is given by:

\[ u(t) = e^{tL}f + \int_0^t e^{(t-\tau)L}g(\tau)d\tau \]  

(3.15)

Letting \(V^s = \sum_{j=1}^{\infty} s^jL_j\), the authors in [36] used the Duhamel’s principle to get the
following:

\[
\begin{align*}
  w(t) &= e^{tL_0} f + \int_0^t e^{(t-\tau)L_0} V^s w(\tau) d\tau \\
  &= e^{tL_0} f + \int_0^t e^{(t-\tau)L_0} V^s e^{\tau L_0} f + \int_0^t \int_0^\tau e^{(t-\tau)L_0} V^s e^{(\tau-\tau_1)L_0} V^s w(\tau_1) d\tau_1 d\tau \\
  &= \ldots
\end{align*}
\]

\[
\begin{align*}
  &= e^{tL_0} f + t \int_{\Sigma_1} e^{\tau_0 L_0} V^s e^{\tau_1 L_0} f d\tau + t^2 \int_{\Sigma_2} e^{\tau_0 L_0} V^s e^{\tau_1 L_0} V^s e^{\tau_2 L_0} f d\tau + \ldots \\
  &+ t^d \int_{\Sigma_d} e^{\tau_0 L_0} V^s e^{\tau_1 L_0} \ldots e^{\tau_d L_0} V^s e^{\tau_{d+1} L_0} f d\tau \\
  &+ t^{d+1} \int_{\Sigma_{d+1}} e^{\tau_0 L_0} V^s e^{\tau_1 L_0} \ldots e^{\tau_d L_0} V^s e^{\tau_{d+1} L} f d\tau
\end{align*}
\]

where

\[
\Sigma_k := \{ \tau = (\tau_0, \tau_1, \ldots, \tau_k) \in \mathbb{R}^{k+1}, \tau_j \geq 0, \Sigma \tau_j = 1 \}
\]

Now the strategy to get the expansion with respect to \( s \) is to collect from each term the like terms: Now because

\[
V^s = L^s - L_0 = \sum_{m=1}^n s^m L_m + L_{n+1}^s,
\]

We can substitute this expression into the last expression in the above iterative formula and collect the terms with the same power of \( s \). i.e. the \( s^0 \) term is simply \( e^{tL_0} f \), The \( s^1 \) term is

\[
\begin{align*}
  &t \int_{\Sigma_1} e^{\tau_0 L_0} L_1 e^{\tau_1 L_0} f d\tau \\
  &= t \int_0^1 e^{t(1-\tau)L_0} L_1 e^{\tau L_0} f d\tau \\
  &= t \int_0^1 e^{t(1-\tau)ad_{L_0}} L_1 e^{t(1-\tau)L_0} e^{\tau L_0} f d\tau \\
  &= t \int_0^1 e^{t(1-\tau)ad_{L_0}} L_1 e^{tL_0} f d\tau \\
  &= t \left( \int_0^1 e^{t(1-\tau)ad_{L_0}d\tau} L_1 \right) e^{tL_0} f \\
  &= t \left( \int_0^1 (L_1 + t(1-\tau)[L_0, L_1]) d\tau \right) e^{tL_0} f \\
  &= t \left( L_1 + \frac{t}{2} [L_0, L_1] \right) e^{tL_0} f
\end{align*}
\]
The $s^2$ term requires more work. It is:
\[
 t \int_0^1 e^{t(1-\tau)L_0} L_2 e^{\tau L_0} f d\tau + t^2 \int_0^1 \int_0^\tau e^{t(1-\tau)L_0} L_1 e^{t(\tau-\tau_1)L_0} L_1 e^{\tau_1 L_0} f d\tau_1 d\tau
\]
The first terms is
\[
 t \left( L_2 + \frac{t}{2}[L_0, L_2] + \frac{t^2}{6}[L_0, [L_0, L_2]] \right) e^{tL_0} f,
\]
because at this time
\[
e^{h_{\text{ad}} L_0} L_2 = L_2 + h[L_0, L_2] + h^2[L_0, [L_0, L_2]].
\]
The second term is
\[
t^2 \int_0^1 \int_0^\tau e^{t(1-\tau)L_0} L_1 e^{t(\tau-\tau_1)L_0} L_1 e^{\tau_1 L_0} f d\tau_1 d\tau
\]
\[
= t^2 \int_0^1 \int_0^\tau e^{t(1-\tau)L_0} L_1 e^{t(1-\tau)L_0} L_1 e^{t(\tau-\tau_1)L_0} L_1 e^{\tau_1 L_0} f d\tau_1 d\tau
\]
\[
= t^2 \int_0^1 \int_0^\tau e^{t(1-\tau)L_0} L_1 e^{t(1-\tau)L_0} L_1 e^{t(\tau-\tau_1)L_0} L_1 e^{\tau_1 L_0} f d\tau_1 d\tau
\]
\[
= t^2 \int_0^1 \int_0^\tau e^{t(1-\tau)L_0} L_1 e^{t(1-\tau)L_0} L_1 e^{t(\tau-\tau_1)L_0} L_1 e^{\tau_1 L_0} f d\tau_1 d\tau
\]
\[
= t^2 \left( \int_0^1 \int_0^\tau e^{t(1-\tau)L_0} L_1 e^{t(1-\tau)L_0} L_1 e^{\tau_1 L_0} f d\tau_1 \right) e^{tL_0} f
\]
\[
= t^2 \left( \int_0^1 \int_0^\tau (L_1 + t(1 - \tau)[L_0, L_1]) (L_1 + t(1 - \tau)[L_0, L_1]) d\tau_1 d\tau \right) e^{tL_0} f
\]
\[
= t^2 \left( \frac{1}{2} L_1^2 + \frac{t}{3} L_1 [L_0, L_1] + \frac{t^2}{6} [L_0, L_1] L_1 + \frac{t^2}{8} [L_0, L_1]^2 \right) e^{tL_0} f
\]
Combining the two terms together, we get for the $s^2$ term:
\[
 t \left( L_2 + \frac{t}{2}[L_0, L_2] + \frac{t^2}{6}[L_0, [L_0, L_2]] \right) e^{tL_0} f
\]
\[
+ t^2 \left( \frac{1}{2} L_1^2 + \frac{t}{3} L_1 [L_0, L_1] + \frac{t^2}{6} [L_0, L_1] L_1 + \frac{t^2}{8} [L_0, L_1]^2 \right) e^{tL_0} f
\]
(3.17)

It’s not that easy to proceed because we will have more and more terms when we collect the coefficients for $s^m$ as $m$ becomes bigger. Now we prove that our perturbative approach is equivalent to the iterative approach shown above in [36]
Theorem 3.3.2. The perturbative expansion approach (2.7) and the iterative approach (3.3) give the same expansion form with respect to $s$.

Proof. A closer look to the iterative approach:

$$w(t) = e^{tL_0}f + \int_0^t e^{(t-\tau)L_0} V^s w(\tau) d\tau$$  \hfill (3.18)

The next step is to insert this expression into the second term. Now suppose the iterative approach gives:

$$w(t) = \sum_{k=0}^{\infty} s^k w_k(t)$$  \hfill (3.19)

We also write explicitly:

$$V^s = \sum_{j=1}^{\infty} s^j L_j$$  \hfill (3.20)

Now we seek how the iterative approach gets $w_{m+1}$, which is got by collecting all terms containing $s^{m+1}$. In other words, it is obtained by collecting all terms containing $s^{m+1}$ in:

$$\int_0^t e^{(t-\tau)L_0} \left( \sum_{j=1}^{\infty} s^j L_j \right) \left( \sum_{k=0}^{\infty} s^k w_k(t) \right)$$  \hfill (3.21)

Simple calculation shows that this term is:

$$\int_0^t e^{(t-\tau)L_0} \left( \sum_{j=1}^{m+1} s^{m+1} L_j w_{m+1-j} \right)$$  \hfill (3.22)

$w_{m+1}$ produced by the iterative approach is given by:

$$w_{m+1} = \int_0^t e^{(t-\tau)L_0} \left( \sum_{j=1}^{m+1} L_j w_{m+1-j} \right)$$  \hfill (3.23)

where $w_0, \ldots, w_m$ are all obtained from the iterative approach. From the previous calculation, it is easy to show that $w_0 = f$. Now by Lemma 2.2.1 and mathematical induction argument, we complete the proof.

3.4 Approximation to $G^L_t(x, y)$ and Error Estimate

On establish the equivalence result of perturbative expansion and iterative expansion in last section, we can follow [36] to establish the convergence result for the Green’s function
expansion. From (3.9), we define
\[ G_m(t, z) := t^{−N/2} \sum_{j=0}^{m} t^j B_j(z, t^{−1/2}(x − z), t^{−1/2}(y − z)) G(t−1/2(x − z), t−1/2(y − z)). \]

While \( B_j \) are explicitly computed, \( G_m(x, y) \) is an approximation to \( G_L(x, y) \) in suitable Sobolev space.

Definition: (Weighted Sobolev space) Let \( \langle x \rangle = (1 + |x|^2)^{1/2} \) \( W^{m,p}_a := W^{m,p}_a(\mathbb{R}^N) \) is the exponentially weighted Sobolev space defined by
\[ W^{m,p}_a(\mathbb{R}^N) := \{ u : \mathbb{R}^N \to \mathbb{C}, \partial^\alpha \langle e^{a(x)} \rangle u(\cdot) \in L^p(\mathbb{R}^N) \}. \]

Then the following Theorem holds:

**Theorem 3.4.1.** Let \( m \in \mathbb{N}, L \in \mathbb{L}_γ \). Define the error term \( E^m_t \) in the approximation of the Green’s function by
\[ e^{tL}f(x) = \int_{\mathbb{R}^N} G^m_t(x, y) f(y) dy + t^{(m+1)/2} E^m_t f(x). \tag{3.24} \]
Then, for any \( f \in W^{m,p}_a(\mathbb{R}^N), a, m \in \mathbb{R}, 1 < p < \infty \), we have
\[ \| E^m_t f \|_{W^{m+k,p}_a} \leq C t^{-k/2} \| f \|_{W^{m,p}_a}, \tag{3.25} \]
for any \( t \in [0, T], k \in \mathbb{N} \), with \( C \) independent of \( t \in [0, T] \).

**Proof.** Refer to Theorem 1.1 in [36], and the equivalence of the iterative expansion and perturbative expansion, we proved the convergence. \( \square \)
Chapter 4  
Approximating the Solutions of Parabolic Equations-Dyson-Taylor and Beyond

In the previous chapter we obtained an approximation to the heat kernel of the parabolic equation (3.3). The solution of the equation is the fundamental solution convoluted with the initial condition. In this chapter, we will convolute the approximated heat kernel with the initial condition, to get the approximate solution. It turns out that no numerical integration will be needed. We will also introduce two other approaches to the approximation problem: the first one is based on a different arrangement of the Taylor expansion of the coefficient, and the other one is more interesting, because no Taylor expansion is needed.

4.1 Approximation for the solution

Recall that we get the approximated heat kernel in the previous chapter:

\[ G_t^L(x, y) = \frac{t^{-N/2}}{\sqrt{(4\pi)^N \det A^0}} \sum_{n=0}^{\infty} t^{n/2} P_n(1) \exp(L_0^z)(x', y') \bigg|_{x'=x+t^{-1/2}(x-z), y'=y+t^{-1/2}(y-z)} \]

Note that the terms in \( P_n(1) \) are differential operators w.r.t. \( x' \), and the coefficients are polynomials of \( (x' - z) \). Note that \( z = z(x, y) \) is admissible, which is defined in last chapter, to guarantee the convergence in certain functional space. Throughout this chapter, we take \( z(x, y) = x \), such that \( P_n(1) \) is independent of \( y \). This choice of \( z \) is admissible. For equation (3.3), we integrate this approximated Green’s function against the initial condition \( u(0, x) = f(x) \)
\[ u(t, x) = \int_{y \in \mathbb{R}^N} G_t^L(x, y) f(y) dy \]
\[ = \int \frac{t^{-N/2}}{\sqrt{(4\pi)^N \det A^0}} \sum_{n=0}^{\infty} t^{n/2} \mathcal{P}_n(1) \exp(L_0^z)(x', y')|_{x'=z+t^{-1/2}(x-z), y'=z+t^{-1/2}(y-z)} f(y) dy \]
\[ = \sum_{n=0}^{\infty} t^{n/2} \mathcal{P}_n(1) \int \frac{t^{-N/2}}{\sqrt{(4\pi)^N \det A^0}} \exp\left( -\left( x' - y' \right)(A^0)^{-1}\left( x' - y' \right)/4 \right)|_{x'=z+t^{-1/2}(x-z), y'=z+t^{-1/2}(y-z)} f(y) dy \]
\[ = \sum_{n=0}^{\infty} t^{n/2} \tilde{\mathcal{P}}_n(1) u^0(t, z + \sqrt{t}(x' - z))|_{x'=z+t^{-1/2}(x-z)} \]
\[ = \sum_{n=0}^{\infty} t^{n/2} \tilde{\mathcal{P}}_n(1) u^0(t, x) \]

Note that the choice \( z(x, y) = x \) is crucial, which enables us to put \( \mathcal{P}_n \) out of the integration. \( \tilde{\mathcal{P}}_n(1) \) is slightly different from \( \mathcal{P}_n(1) \), where each terms in the \( \mathcal{P}_n(1) \) which is \( p_{i,j}^n(x - z)^i \frac{\partial}{\partial x^j} \) turns to be \( p_{i,j}^n(x - z)^i \sqrt{t} \frac{\partial}{\partial x^j} \) in \( \tilde{\mathcal{P}}_n(1) \), because of the substitution \( x' = z + t^{-1/2}(x - z) \). Also note that since we have already chosen \( z = x \), \( \tilde{\mathcal{P}}_n(1) \) is actually only consists of terms in the form of \( p_{i,j}^n \sqrt{t} \frac{\partial}{\partial x^j} \), where \( p_{i,j}^n \) comes from the commutator computation of \( \mathcal{P}_n(1) \), and only depends on the coefficients \( a(x), b(x), c(x) \) and there derivatives, at the point \( x \). For one dimensional, we have explicitly computed terms in \( \tilde{\mathcal{P}}_n(1) \), and list for the first few values of \( n \):

\[ \tilde{\mathcal{P}}_1(1) = b(x) \sqrt{t} \frac{\partial}{\partial x} + a(x) a'(x) \sqrt{t^3} \frac{\partial^3}{\partial x^3} \]
\[ \tilde{\mathcal{P}}_2(1) = c(x) + (b(x))^2 + a'(x) b(x) + a a''(x) + 2ab'(x))/2\sqrt{t} \frac{\partial}{\partial x} + (a(2a'(x))^2 + 2b(x)a'(x))/3 + (a(x)a'(x))^2)/3 + (2a(x)^2a'(x))/3 + (aa'(x)b(x))/3\sqrt{t} \frac{\partial^4}{\partial x^4} \]
\[ = (a^2(x)a'^2(x))\sqrt{t^6} \frac{\partial^6}{\partial x^6} \]

Higher order \( \tilde{\mathcal{P}}_n \) are available by the symbolic computation codes. Also note that \( u^0(t, x) \) solved the constant coefficient PDE:

\[ \partial_t u^0(t, x) = a(z) \partial_x^2 u^0(t, x) \]  \hspace{1cm} (4.2)

with the same initial condition with (3.3). \( z = x \) is freezes, and \( x \) is the point we are interested in. So be careful that this expansion solution is only correct for \( x \), since \( \tilde{\mathcal{P}}_n(1) \) depends on \( x \).
A simple example: for the 1 dimensional constant coefficient equation:

$$\partial_t u = a\partial_{xx}^2 u + b\partial_x u$$

(4.3)

with initial condition $u(0, x) = f(x)$, we have: $\hat{P}_n(1) = (b\sqrt{t})^n\partial_x^n$ and

$$u(t, x) = \sum_{n=0}^{\infty} \hat{P}_n(1)u^0(t, x)$$

$$= \sum_{n=0}^{\infty} (bt\partial_x)^n u^0(t, x)$$

$$= \exp(bt\partial_x)u^0(t, x)$$

$$= u^0(t, x + bt)$$

where $u^0(t, x)$ solves:

$$\partial_t u^0 = a\partial_{xx}^2 u^0$$

Such that we see we recovered the true solution, by collecting infinitely many terms of the expansion.

From this example, we also see space to improve the approximation: even for constant coefficient case, the Dyson-Taylor commutator method needs to collect infinitely main terms to recover the true solution. On the other hand, one may want to connect the true solution of the (3.3) directly to the solution to the constant coefficient case:

$$\partial_t u = a(z)\partial_{xx}^2 u + b(z)\partial_x u + c(z)u$$

(4.4)

We came up with new approaches for the approximation of the heat kernel, by a new scheme of expansion. We also prove that we recover the Dyson-Taylor commutator method, in case we arrange the expansion in certain order. On the other hand, $z(x, y)$ does not have to be equal to $x$, which is crucial for approximation to CEV and SABR models in later chapters.
4.2 Different approaches for the approximation of the heat kernel

Recall we are solving the parabolic equation:

\[ \partial_t u = a(x)\partial_x^2 u + b(x)\partial_x u + c(x)u \quad (4.5) \]

Now fix \( z \in \mathbb{R}^N \), we rewrite the above as:

\[ \partial_t u = a(z)\partial_x^2 u + b(z)\partial_x u + c(z)u + (a(x) - a(z))\partial_x^2 u + (b(x) - b(z))\partial_x u + (c(x) - c(z))u \quad (4.6) \]

and Taylor expand the difference of the functions to get:

\[ \partial_t u = a(z)\partial_x^2 u + b(z)\partial_x u + c(z)u + \sum_{j=1}^{\infty} \frac{(x - z)^j}{j!} (a^{(j)}(z)\partial_x^2 u + b^{(j)}(z)\partial_x u + c^{(j)}(z)u) \quad (4.7) \]

At this stage, it is up to us to decide the order of the operators: Two cases are of primary interest:

Case 1:

Letting

\[ L_j = \frac{(x - z)^j}{j!} a^{(j)}(z)\partial_x^2 u + \frac{(x - z)^{j-1}}{(j-1)!} b^{(j-1)}(z)\partial_x u + \frac{(x - z)^{j-2}}{(j-2)!} c^{(j-2)}(z)u \quad (4.8) \]

where we treat \( b^{(j)} = c^{(j)} = 0 \) if \( j < 0 \). Recall Theorem 2.2.1 we have the following:

\[ u(t, x) = \sum_{j=0}^{\infty} P_n(t)u^0(t, x) \quad (4.9) \]

where \( P_n \) are given inductively by Lemma 2.2.1, or explicitly by Theorem 2.2.3. \( L_j \) are given by (4.8), and \( u^0(t, x) \) solves the equation:

\[ \partial_t u^0 = a(z)\partial_x^2 u^0 \quad (4.10) \]

We have the following theorem:
Theorem 4.2.1. The approximating series we got from above is exactly the Dyson-Taylor commutator expansion series.

Proof. Carefully compare the relationship between parabolic rescaling and the direct expansion.

Case 2: We are more interested in this second case, where the solution is directly connected with the constant coefficient solution. Furthermore, the expansion series recovers the true solution at finite cutoff (actually the leading term only), in the case the original parabolic PDE has constant coefficients. Letting:

\[ L_j = \frac{(x-z)^j}{j!}(a^{(j)}(z)\partial_x^2 u + b^{(j)}(z)\partial_x u + c^{(j)}(z)u) \]  

where \( a^{(j)} \) is the \( j \)-th derivative of \( a \). Again, recall Theorem 2.2.1 for the small parameter expansion:

\[ u(t, x) = \sum_{j=0}^{\infty} P_n(t)u^0(t, x) \]  

where \( P_n \) are given inductively by Lemma 2.2.1, or explicitly by Theorem 2.2.3. \( L_j \) are given by (4.11), and \( u^0(t, x) \) solves the equation:

\[ \partial_t u^0 = a(z)\partial_x^2 u^0 + b(z)\partial_x u^0 + c(z)u^0 \]  

We have the following theorem:

Theorem 4.2.2. The series expansion (4.12) converges in a suitable functional space to the true solution, for small time \( t \).

This suitable functional space is the weighted Sobolev space, which will be introduced in next section.

Again, for the one dimensional case, we list \( P_n(t) \) for the first few \( n \). More expressions can be found by symbolic computation:

\[ P_1(t) = t^2(bc' + (2ac' + bb')\partial + (2ab' + a'b)\partial^2 + 2aa'\partial^3)/2; \]
\[ P_2(t) = t^2(ac'' + ab''\partial_x + aa''\partial_x^2)/2 + t^3... + t^4...; \]

one can still proceed to prove that the highest order term of derivative in \( P_j \) is \( 3j \), and the order of time variable is between 2 and 2j.
In the following, we will provide a time-rescaled version of (3.3), and this is equivalent to the original problem. However, the numerical computation terms out to be simpler in the time-rescaled case: For fixed \( z \), we let \( v(t, x) = u(\frac{t}{2a(z)}, x) \), then \( v(t, x) \) solves the equation:

\[
\partial_t v = \frac{a(x)}{2a(z)} \partial_x^2 v + \frac{b(x)}{2a(z)} \partial_x v + \frac{c(x)}{2a(z)} v
\]

(4.14)

and such that

\[
v(t, x) = \sum_{j=0}^{\infty} \hat{P}_n(t) v^0(t, x)
\]

(4.15)

where \( \hat{P}_n \) are given inductively by Lemma 2.2.1, or explicitly by Theorem 2.2.3. \( \hat{L}_j \) are given by:

\[
\hat{L}_j = \frac{(x-z)^j}{2a(z)j!} (a^{(j)}(z) \partial_x^2 u + b^{(j)}(z) \partial_x u + c^{(j)}(z) u),
\]

(4.16)

and \( v^0(t, x) \) solves the equation:

\[
\partial_t v^0 = \frac{1}{2} \partial_x^2 v^0 + \frac{b(z)}{2a(z)} \partial_x v^0 + \frac{c(z)}{2a(z)} v^0
\]

(4.17)

Recover \( u(t, x) \), we have

\[
u(t, x) = \sum_{j=0}^{\infty} \hat{P}_n(2a(z)t) u^0(t, x)
\]

(4.18)

We will use this time scaled version of the expansion in most of the following chapters.

4.3 A splitting method where Taylor expansion is not needed

In previous section, we introduced a new ordering of operators which are different from the original Dyson-Taylor method. In this section, we will introduce a method of computing the approximate solution of (3.3), which does not involved Taylor expansion of the coefficients. This method is new as far as I know. It turns out that the computation of the implied volatility will use both the new ordering of Dyson-Taylor method, and the method introduced in this section.

Recall that we are interested in the problem:

\[
Lu(x) := \sum_{i,j=1}^{N} a_{ij} \partial_i \partial_j u(x) + \sum_{k=1}^{N} b_k(x) \partial_k u(x) + c(x) u(x),
\]

(4.19)
where \( x = (x_1, \ldots, x_N) \in \mathbb{R}^N \), \( \partial_k := \frac{\partial}{\partial x_k} \), and the coefficients \( a_{ij}, b_i \) and \( c \) and all their derivatives are assumed to be smooth and uniformly bounded.

For the initial value problem:

\[
\begin{align*}
\partial_t u(t, x) - Lu(t, x) &= g(t, x) \quad \text{in } (0, \infty) \times \mathbb{R}^N, \\
u(0, x) &= f(x) \quad \text{on } \{0\} \times \mathbb{R}^N,
\end{align*}
\]

(4.20)

for \( u, f, \) and \( g \) in suitable function spaces. In view of Duhamel’s principle, we may assume \( g = 0. \)

The solution can be written formally by the semi-group expression:

\[
u(t, x) = e^{tL}f
\]

(4.21)

### 4.3.1 An Example: Matrix Case

Mathematical models of many physical, biological and economic processes involve systems of linear, constant coefficient ordinary differential equations (ODE)

\[
\dot{x}(t) = Ax(t).
\]

Here \( A \) is a given, fixed, real or complex \( n \) by \( n \) matrix. A solution vector \( x(t) \) is sought which satisfies an initial condition

\[
x(0) = x_0.
\]

The solution of this problem is

\[
x(t) = e^{tA}x_0.
\]

where the matrix exponential is defined by the convergent power series:

\[
e^{tA} = I + tA + \frac{t^3A^2}{2!} + \cdots
\]

(4.22)

A good survey of matrix computation is [67], where 19 different methods were introduced. Methods involving approximation theory, differential equations, the matrix eigenvalues, and the matrix characteristic polynomial have been proposed. We also refer to [68] for reader who are interested in general matrix computations.
4.3.2 Operator splitting and a new approach to compute matrix exponential

To compute $e^A$ for a matrix $A \in \mathbb{M}_{n \times n}$, we decompose it into $A = X + Y$, and have the following theorem:

**Theorem 4.3.1.**

$$e^{X+Y} = \left(e^{ad_{X+Y}I_{n \times n}}\right) e^X$$

(4.23)

where $I_{n \times n}$ is the identity matrix. $e^{ad_{X+Y}}$ is defined as a linear operator for matrix:

$$e^{ad_{X+Y}}Z = Z + (adX + Y)Z + \frac{(adX + Y)^2 Z}{2!} + \ldots$$

$$= Z + ([X, Z] + YZ) + \frac{[X, [X, Z] + YZ] + Y([X, Z] + YZ)}{2!} + \ldots$$

Recall the adjoint operator for a matrix $X$ is defined by:

$$ad_{X}Y = [X, Y] = XY - YX.$$

**Proof.** We solve the following ODE:

$$C'(t) = (X + Y)C(t)$$

(4.24)

with $C(0) = I_{n \times n}$. The solution is obviously:

$$C(t) = e^{t(X+Y)}$$

(4.25)

On the other hand, we claim that the solution of (4.24) can be expressed:

$$D(t) = \left(e^{t(ad_{X+Y})I_{n \times n}}\right)e^{tX}$$

(4.26)

Directly computation shows:

$$D'(t) = (ad_{X} + Y) \left(e^{t(ad_{X+Y})I_{n \times n}}\right)e^{tX} + \left(e^{t(ad_{X+Y})I_{n \times n}}\right)e^{tX}X$$

$$= (ad_{X} + Y)D(t) + D(t)X$$

$$= [X, D(t)] + YD(t) + D(t)X$$

$$= XD(t) - D(t)X + YD(t) + D(t)X$$
\((X + Y) D(t)\)

Notice that \(D(0) = I_{n \times n}\). By the uniqueness of the solution to (4.24):

\[ e^{t(X+Y)} = \left( e^{t(\text{ad}_X+Y)} I_{n \times n} \right) e^{tX}. \]

The theorem is proved by letting \(t = 1\).

Notice that we used the fact that \(e^{tL}\) is differentiable for any bounded operator \(L\), and the derivative is \(Le^{tL}\), which is the case for \(L\) being a square matrix.

### 4.3.3 Differential operator case

Let us go back to the original problem (3.3). The solution \(u(t, \cdot)\) can be expressed in the following semi-group expression by (4.21). However, this expression in general does not allow a convergent power series similar to (4.22), unless the initial condition \(f\) is sufficiently smooth, which is not always the case. For example, the initial condition for European call option price has discontinuous first order derivative. Thus, this formula does not provide us more information than the original problem (3.3).

We split \(L\) into

\[ L = L_0 + V, \]

where \(L_0 \in \mathcal{L}_{\gamma}\), and \(V\) is ‘small’ in the sense which will be explained later. Then we write down a formula which is similar to (4.23)

\[ e^{tL} = e^{t(L_0+V)} \sim \left( e^{t(\text{ad}_{L_0}+V)} Id \right) e^{tL_0} \quad (4.27) \]

Notice that we did not use ‘\(=\)’ between last two expressions, and we will explain why this is actually an asymptotic expansion, for small time \(t\).

### 4.3.4 Weighted Sobolev space

In order to explain the meaning of (4.27), we need to introduce suitable functional spaces, and it turns out that exponentially weighted Sobolev space is a good choice, since we are interested in considering the initial value problem (3.3) in the largest-possible space of initial data \(f\) that includes the typical initial conditions that arise in applications (like option pricing problems), and where uniqueness holds. Please see [35,36] for further details.
We introduce the weight \( \langle x \rangle = (1 + |x|^2)^{1/2} \). Define \( W^m_p(R^N) \) for \( m \in \mathbb{Z}_+, a \in \mathbb{R}, 1 < p < \infty \), by
\[
W^r_p(R^N) := e^{-a(x)}W^r_p(R^N)
\]
for \( r \in \mathbb{Z}_+ \). With norm
\[
||u||_{W^r_p}^p = ||e^{a(x)}u||_{W^r_p}^p = \sum_{|\alpha| < r} ||\partial^\alpha_x (e^{a(x)}u(x))||_{L^p}^p
\]
\( W^m_p(R^N) \) is a Banach space. We observe that \( W^m_p(R^N) = W^m_p(R^N) \).

A crucial observation is that, for any \( L \in \mathbb{L}_\gamma \) and any \( a \in \mathbb{R} \), the operators \( L_a := e^{a(x)}Le^{-a(x)} \) are also in \( \mathbb{L}_\gamma \). They moreover define a bounded family in \( \mathbb{L}_\gamma \) if \( a \) is in a bounded set. Since proving a result for \( L \) acting between weighted Sobolev spaces \( W^{s,p}_a \) is the same as proving the corresponding result for \( L_a \) acting between the Sobolev spaces \( W^{s,p}_a = W^{s,p}_0 \), we may assume that \( a = 0 \). In particular, \( L : W^{s+2,p}_a \to W^{s,p}_a \) is well defined and continuous for any \( a \), since this is true for \( a = 0 \).

We will use the following proposition, which is Proposition 2.7 on [36]:

**Proposition 4.3.2.** Let \( L \) be an arbitrary element in \( \mathbb{L}_\gamma \). We have \( e^{tL}W^{s+p}_a \subset W^{r+p}_a \) for all \( r, s, a \in \mathbb{R}, 1 < p < \infty, \) and \( t > 0 \). Let \( r \geq s \), then
\[
||u(t)||_{W^{r,p}_a} \leq Ct^{(s-t)/2}||f||_{W^{s,p}_a}, \quad t \in (0, 1].
\]
The constant \( C \) above is independent of \( r, s, a, p, \) and \( L \), as long as they belong to bounded sets.

We need some more propositions for further use. We prove the following:

**Proposition 4.3.3.** Let \( L_0 \in \mathbb{L}_\gamma \), \( V \) be a first order differential operator with smooth and bounded coefficients for any order. Then for any non-negative integer \( k \), we have:
\[
|| \left( (ad_{L_0} + V)^k Id \right) f||_{W^{r,p}_a} \leq C||f||_{W^{r+k,p}_a}
\]
for all \( a, m, r, p \in \mathbb{R} \), and all smooth \( f \). Furthermore, \( C \) is independent of \( f \).

**Proof.** We notice that \( (ad_{L_0} + V)^k Id := (ad_{L_0} + V) \left( (ad_{L_0} + V)^{k-1} Id \right) \) has finitely many terms, and the highest order of differential being \( k \), and all coefficients being bounded, which
could easily be proved by induction. Then the proposition follows because the identity embeddings $W^{m+1,p}_a \to W^{m,p}_a$ are continuous for all $m \in \mathbb{Z}_+, 1 < p < \infty, a \in \mathbb{R}$.

**4.3.5 Proof of asymptotic expression**

Recall (4.27). For the problem (3.3) with $g(t, x) = 0$, we write

$$u_N(t) = \left( \sum_{k=0}^{N} \frac{t^k(ad_{L_0} + V)^k}{k!}Id \right) e^{tL_0} f$$

(4.30)

Notice that

$$(ad_{L_0} + V)^m L' := (ad_{L_0} + V) \left( (ad_{L_0} + V)^{m-1} L' \right)$$

for $m \geq 1$. We also notice that for $f \in W^{s,p}_a$ for some $a, s, p \in \mathbb{R}$, $e^{tL_0} f \in W^{r,p}_a$ for any $r \in \mathbb{R}$, for $t > 0$. In other words, $e^{tL_0} f$ is smooth enough, such that $v_N$ is well defined in classical sense.

It is straightforward to verify that $t \to v_N(t)$ is smooth for $0 < t < \infty$, and direct computation shows:

$$\partial_t u_N(t) = \left( \sum_{k=0}^{N} \frac{kt^{k-1}(ad_{L_0} + V)^k}{k!}Id \right) e^{tL_0} f + \left( \sum_{k=0}^{N} \frac{tk(ad_{L_0} + V)^k}{k!}Id \right) L_0 e^{tL_0} f$$

$$= (ad_{L_0} + V) \left( \sum_{k=0}^{N-1} \frac{tk(ad_{L_0} + V)^k}{k!}Id \right) e^{tL_0} f + \left( \sum_{k=0}^{N} \frac{tk(ad_{L_0} + V)^k}{k!}Id \right) L_0 e^{tL_0} f$$

$$= (L_0 + V) \left( \sum_{k=0}^{N} \frac{tk(ad_{L_0} + V)^k}{k!}Id \right) e^{tL_0} f - \left( \frac{t^N(ad_{L_0} + V)^{N+1}}{N!}Id \right) e^{tL_0} f$$

$$= Lu_N - g_N(t, x)$$

where

$$g_N(t, x) = \left( \frac{t^N(ad_{L_0} + V)^{N+1}}{N!}Id \right) e^{tL_0} f$$

Letting $u = e^{tL} f$ be the solution to (3.3), and

$$w_N = u - u_N,$$

(4.31)

we have:

$$\partial_t w_N = \partial_t u - \partial_t u_N$$

51
\[ = Lu - Lu_N + g_N \]
\[ = Lw_N + g_N \]

with initial condition \( w_N(0) = 0 \). We apply Duhamel’s principle to get:

\[
w_N = \int_0^t e^{(t-\tau)L} g_N(\tau) d\tau \\
= \int_0^t e^{(t-\tau)L} \left( \frac{\tau^N(ad_{L_0} + V)^{N+1}}{N!} Id \right) e^{\tau L_0} f d\tau
\]  

(4.32)

Now we use Proposition 4.3.2 and 4.3.3, to carefully control the norm of \( w_N \) in suitable exponentially weighted Sobolev space:

\[
\|e^{(t-\tau)L} ((ad_{L_0} + V)^{N+1}Id) e^{\tau L_0} f \|_{W^{r,p}_a} \\
\leq C \| ((ad_{L_0} + V)^{N+1}Id) e^{\tau L_0} f \|_{W^{r,p}_a} \quad \text{by Proposition 4.3.2}
\]

\[
\leq C \|e^{\tau L_0} f \|_{W^{r+N+1,p}_a} \quad \text{by Proposition 4.3.3}
\]

\[
\leq C \tau^{(m-r-N-1)/2} \|f\|_{W^{m,p}_a} \quad \text{by Proposition 4.3.2}
\]

Return to (4.32), we obtain:

\[
\|w_N\|_{W^{r,p}_a} \leq \int_0^t \frac{\tau^N}{N!} \|e^{(t-\tau)L} ((ad_{L_0} + V)^{N+1}Id) e^{\tau L_0} f \|_{W^{r,p}_a} d\tau \\
\leq \int_0^t \frac{\tau^N}{N!} C \tau^{(m-r-N-1)/2} \|f\|_{W^{m,p}_a} d\tau \\
\leq C t^{(N+1)/2} \|f\|_{W^{m,p}_a}
\]  

(4.33)

Here \( C \) depends on \( N, m, r, L_0, V \), but is independent of \( t \) and \( f \), for \( 0 < t < T \), for \( 0 < T < \infty \).

We have the following Theorem:

**Theorem 4.3.4.** Let \( L \in \mathbb{L}_\gamma \) be a strongly elliptic operator on \( \mathbb{R}^N \), and \( L_0 \in \mathbb{L}_\gamma \), \( V \) be a first order differential operator such that all its coefficients are bounded for any high order, and that \( L = L_0 + V \). Then we have the asymptotic expression of the solution to (3.3) by:

\[
u(t, \cdot) = e^{tL} f \sim (e^{t(ad_{L_0} + V)} Id) e^{tL_0} f
\]  

(4.34)
In the sense that if we define $\mathcal{E}_{t,N}$ by:

$$u(t, \cdot) = u_N(t, \cdot) + t^{(N+1)/2} \mathcal{E}_{t,N} f$$

where $u_N$ is given by (4.30). Then, for any $f \in W^{m,p}_a(\mathbb{R}^N), a \in \mathbb{R}, m \geq 0, 1 < p < \infty$, we have

$$||\mathcal{E}_{t,N} f||_{W^{m+k,p}_a} \leq Ct^{-k/2}||f||_{W^{m,p}_a},$$

for any $t \in [0,T], 0 < T < \infty, k \in \mathbb{Z}_+, with C independent of t \in [0,T], and f$.

Proof. By (4.31) and (4.33). 

We will show an application of this new splitting method in next chapter where the transition kernel is computed for the SABR model.
Chapter 5  
Two and Higher Dimensional Approximation: Heston, SABR models

We have already explored the perturbative expansion for one dimensional parabolic equations, with applications to CEV model. Next, we will extend our approximation framework to higher dimensions. Keep in mind that we eventually want to approximate the solution for stochastic volatility models, such as Heston model and SABR model. In the first two sections, we seek the option price and marginal transition kernel on the price variable. However, in some cases, such as likelihood-based inference, the full transition kernel is also of importance, and we will treat this situation in the last section of this chapter by exploring the transition kernel for the SABR model.

5.1 Higher dimensions: when the initial condition only depends on the first coordinate

Option pricing is an interesting special case because even in a higher dimensional model, such as those with stochastic volatility, the initial condition depends on only the one variable: the stock price

For the initial value problem

\[
\begin{aligned}
\begin{cases}
\partial_t u(t, x) - Lu(t, x) = 0 & \quad \text{in } (0, \infty) \times \mathbb{R}^N \\
u(0, x) = f(x) & \quad \text{on } \{0\} \times \mathbb{R}^N,
\end{cases}
\end{aligned}
\]  

(5.1)
where
\[
Lu(x) := \sum_{i,j=1}^{N} a_{ij}(x_1, \tilde{x}) \partial_i \partial_j u(x) + \sum_{k=1}^{N} b_k(x_1, \tilde{x}) \partial_k u(x) + c(x_1, \tilde{x}) u(x),
\] (5.2)

where \( \tilde{x} = (x_2, \ldots, x_N) \)

Given the initial condition
\[
u(0, x) = f(x_1, \tilde{x}) = g(x_1),
\] (5.3)

which depends only on the first variable \( x_1 \), We have the following theorem:

**Theorem 5.1.1.** The solution of the problem (5.1) is given by:
\[
u(t, x) = \sum_{n=0}^{\infty} \tilde{P}_n(t) v^0(x, t)
\] (5.4)

where \( v^0 \) solves the one dimensional constant coefficients problem:
\[
\partial_t v^0(t, x) = a_{11}(z_1, \tilde{z}) \partial_x^2 v^0(x) + b_1(z_1, \tilde{z}) \partial_x v^0 + c(z_1, \tilde{z}) v^0
\] (5.5)

with initial condition:
\[
v^0(0, x) = f(x) = g(x_1)
\] (5.6)

**Proof.** We need to prove that \( v^0 \) solves the PDE:
\[
\partial_t v^0 = L_0 v^0,
\] (5.7)

with initial condition
\[
v^0(0, x) = g(x_1).
\]

where
\[
L_0 = \sum_{i,j=1}^{N} a_{ij}(z_1, \tilde{z}) \partial_i \partial_j + \sum_{k=1}^{N} b_k(z_1, \tilde{z}) \partial_k + c(z_1, \tilde{z}).
\]

We notice that \( v^0 \) is independent of \( \tilde{x} = (x_2, \ldots, x_n) \), such that:
\[
\partial_t v^0 = a_{11}(z_1, \tilde{z}) \partial_x^2 v^0(x) + b_1(z_1, \tilde{z}) \partial_x v^0 + c(z_1, \tilde{z}) v^0 = L_0 v^0.
\]

Meanwhile, since the initial condition agrees, by the uniqueness of the solution to (5.7), we proved this theorem.\( \square \)
Notice that this result tells us that if the initial condition is independent of the other coordinates except for the first one, we can always find the solution by perturbing the solution of a 1D problem. Furthermore, all the terms in \( \tilde{P}_n(1) \) are in the form of \( p_{i_1}^{j_1} \cdots i_n^{j_n} \Pi_{j=1}^n (x_j - z_j)^{i_j} \partial_{x_1}^{j_1} \) if we use the Dyson Taylor approach, which makes the approximating formula extremely clean. The dependence of the solution on variables \( \tilde{x} = (x_2, \cdots, x_n) \) is from 1) the ‘frozen point’ \( \tilde{z} \), and 2) the computation of \( \tilde{p}_n \).

Recall that we have developed the theory of approximating higher dimensional parabolic equations in chapter 4, especially, when the initial condition only depends on the first coordinate, we can approximate the solution based on the solution of 1D problem. This is perfectly applicable to stochastic volatility models, since the payoff function only depends on the stock price. Volatility plays its role by influencing the stock price process (indirect influence).

### 5.1.1 Short time expansion for the Heston’s model

Recall the Heston’s model, after the log change of the stock price variable, we have

\[
\partial_t u(t; x, V) = .5V \partial_x^2 u + \rho \sigma_v V \partial_{xv} u + .5\sigma_v^2 V \partial_v^2 u + (r - .5V) \partial_x u + \kappa (\theta - V) \partial_v - ru
\]  

(5.8)

Now we follow the approximating formula (5.4):

\[
u(t, x) = \sum_{n=0}^{\infty} \tilde{P}_n(t) u^0(x, t)
\]  

(5.9)

where \( u^0 \) solves the 1d constant coefficients problem:

\[
\partial_t u^0(t, x) = .5V_0 \partial_x^2 u^0(x) + (r - .5V_0) \partial_x u^0 - ru^0
\]  

(5.10)

with initial condition:

\[
u^0(0, x) = \max(e^x - K, 0)
\]  

(5.11)

We have \( u^0(t, \log(S_0)) = C_{BS}(S_0, r, t, \sqrt{V_0}) \), which provides us as the zero-th order approximation for the Heston Model. We compute \( \tilde{P}_u(t) \) by symbolic computation code, and only collect those terms without \( (x - z_x), (V - z_V) \), since we choose \( z_x = x_0 \) and \( z_V = V_0 \).

Again, we will use the time scaled version, such that: \( T = V_0 t \), and

\[
L_0 = .5\partial_x^2 u + \rho \sigma_v \partial_{xv} u + .5\sigma_v^2 \partial_v^2 u + (r/V_0 - .5) \partial_x + \kappa (\theta/V_0 - 1) \partial_v - r/V_0 u
\]
Figure 5.1: Heston model European call price, with $t = .25, \sigma_V = .32, \theta = .36, \rho = -.01, V_0 = .24, S_0 = 100, r = .05$, and various strikes

$$L_1 = .5(V/V_0 - 1)\partial_x^2 u + (V/V_0 - 1)\rho \sigma_v \partial_x \partial_v u + .5(V/V_0 - 1)\sigma_v^2 \partial_v^2 u - .5(V/V_0 - 1)\partial_x - \kappa(V/V_0 - 1)\partial_v$$

$L_j = 0$, for $j \geq 2$

We present the numerical experiment in Figure 5.1b for the approximating solution of Heston model, for the European call option price. From the figure, we notice that it is necessary to have higher order approximations, in order to obtain acceptable price error, as well as the percentage error, especially for deep out of money options.

### 5.1.2 Short time expansion for the SABR model

The SABR model is widely used in industry. In this section, we will explore the closed form approximation to the extended SABR model, which is called $\lambda$-SABR [21]:

$$\begin{cases}
    dF_t = \sigma_t F_t^\beta dW_t \\
    d\sigma_t = \kappa(\theta - \sigma_t)dt + \alpha \sigma_t dZ_t
\end{cases}$$


Here $F_t$ is the forward price, and $dW_t dZ_t = \rho dt$. Notice that when $\kappa = 0$, we recover the original version of SABR model.

The general methodology for treating this model is essentially the same as the Heston model. We use the log change of the stock price variable to get:

$$\partial_t u(t; x, V) = 0.5 V \partial^2_x u + \rho \sigma_V V \partial^2_{xv} u + 0.5 \sigma^2 V \partial^2_v u + (r - 0.5 V) \partial_x u + \kappa (\theta - V) \partial_v - ru$$  \hspace{1cm} (5.12)\\

Now we follow the approximating formula (5.13).

$$u(t, x) = \sum_{n=0}^{\infty} \tilde{P}_n(t) u^0(x, t)$$  \hspace{1cm} (5.13)\\

where $u^0$ solves the 1d constant coefficients problem:

$$\partial_t u^0(t, x) = 0.5 V_0 \partial^2_x u^0(x) + (r - 0.5 V_0) \partial_x u^0 - ru^0$$  \hspace{1cm} (5.14)\\

with initial condition:

$$u^0(0, x) = \max(e^x - K, 0).$$  \hspace{1cm} (5.15)\\

Now we have $u^0(t, \log(S_0)) = C_{BS}(S_0, r, t, \sqrt{V_0})$, which provides us as the zero-th order approximation for the Heston Model. We compute $\tilde{P}_n(t)$ by symbolic computation code, and only collect those terms without $(x - x_0), (V - V_0)$, as well as $\partial_v$, since we are looking for the value of $u$ exactly at $(t, x_0, V_0)$, and since $u^0$ is independent of $V$. (Although $u_0$ depends on $V_0$, in the sense that $V_0$ is the ‘frozen’ point.) Letting $m = \beta - 1$, $s = 2(\beta - 1)$ and $V_{eff} = \sigma_0^2 S_0^\alpha$

again, one use the time scaled version, such that: $T = V_0 t$, and

$$L_0 = 0.5(\partial^2_x - \partial_x) + \frac{\rho \alpha}{\sigma_0 S_0} \partial^2_{x\sigma} + \frac{\alpha^2}{2 \sigma_0^2} \partial^2_\sigma + \frac{\kappa (\theta - \sigma_0)}{V_{eff}} \partial_\sigma$$

$$L_1 = 0.5 \left( \frac{2(\sigma - \sigma_0)}{\sigma_0} + s(x - x_0) \right) \left( \partial^2_x - \partial_x \right) + \frac{\rho \alpha}{\sigma_0 S_0^m} \left( \frac{2(\sigma - \sigma_0)}{\sigma_0} + m(x - x_0) \right) \partial^2_{x\sigma} + \frac{\alpha^2 (\sigma - \sigma_0)}{\sigma_0 S_0^e} \partial^2_\sigma - \frac{\kappa (\sigma - \sigma_0)}{V_{eff}} \partial_\sigma$$

$$L_2 = 0.5 \left( \frac{(\sigma - \sigma_0)^2}{\sigma_0^2} + 2 s(x - x_0)(\sigma - \sigma_0) + 5 s^2 (x - x_0)^2 \right) \left( \partial^2_x - \partial_x \right) + \frac{\rho \alpha}{\sigma_0 S_0^m} \left( \frac{(\sigma - \sigma_0)^2}{\sigma_0^2} + 2 m(x - x_0)(\sigma - \sigma_0) \right)$$

$$+ \frac{5 m^2 (x - x_0)^2}{\sigma_0} \partial^2_{x\sigma} + \frac{\alpha^2 (\sigma - \sigma_0)^2}{2 V_{eff}} \partial^2_\sigma$$

$$L_j = 0.5 \left( \frac{(\sigma - \sigma_0)^2}{(j - 2)! \sigma_0^2} s^{j-2} (x - x_0)^{-j-2} \right) + \frac{\rho \alpha}{\sigma_0 S_0^m} \left( \frac{(\sigma - \sigma_0)}{(j - 1)! \sigma_0} s^{j-1} (x - x_0)^{j-1} + \frac{1}{j!} s^j (x - x_0)^j \right) \left( \partial^2_x - \partial_x \right)$$
We can see that even for deep out of the money calls, the 7-th order approximation has the relative pricing error within 1%. This again, illustrate the necessity of obtaining higher order approximations. The absolute price difference for lower strikes might be from the error of the Monte Carlo simulations.

5.2 Transition kernel for the SABR model

In this section, we explore the transition kernel

\[ p(t; F_T, \sigma_T|F_0, \sigma_0) \]

of the SABR model. In order to do so, we need to compute the full operators \( P_n \), in which the tricks in the previous two sections do not apply. In this section, I will also show how to use the second way of operator splitting skills, which was introduced in the last section of
Chapter 4.

Recall the SABR PDE:

\[
\begin{align*}
\frac{\partial_t u(t,x) - L u(t,x)}{u(0,x) = f(x) := \delta_{F_T,\sigma_T}} &= 0 \quad \text{in } (0, \infty) \times \mathbb{R}^N \\
&\quad \text{on } \{0\} \times \mathbb{R}^N,
\end{align*}
\]

(5.16)

where \(\delta_{F,\sigma}\) is Dirac’s \(\delta\) function, \(x = (F,\sigma)\), and

\[
L = \frac{\sigma^2 F^{2\beta}}{2} \partial^2_F + \rho\sigma^2 F \beta \partial_{F\sigma} + \frac{\alpha^2 \sigma^2}{2} \partial^2_{\sigma}.
\]

We are interested in the case where \(x = (F_0,\sigma_0)\). Notice that we write the PDE as Kolmogorov’s backward equation, such that we keep the form of the PDE simple (compared with the forward equation). Now we split \(L\) into

\[
L = L_0 + V,
\]

where

\[
L_0 = \frac{\sigma^2 z^2 (\beta - 1)}{2} F^2 \partial^2_F + \rho \sigma z^2 \beta F \partial_{F\sigma} + \frac{\alpha^2 \sigma^2}{2} \partial^2_{\sigma}
\]

and

\[
V = L - L_0,
\]

where \(F_z\) and \(\sigma_z\) are two ‘frozen’ constants. Here we split \(L\) in the above way because \(e^{tL_0}f\) is explicit:

\[
p_{L_0}(t; F_t, \sigma_t | F_0, \sigma_0) := e^{tL_0}f = \frac{e^{-W_F^{\frac{1}{2}} - W_H \sqrt{\frac{1 - \rho^2}{2}}}}{2\pi t F_0 \sigma_0 A \alpha \sqrt{1 - \rho^2}}
\]

(5.17)

where

\[
W_F = \frac{\log F_t - \log F_0 + \sigma^2 A^2 t / 2}{\sigma_z A},
\]

\[
W_H = \frac{\log \sigma_t - \log \sigma_0 + \alpha^2 t / 2 - \rho \sigma W_F}{\alpha \sqrt{1 - \rho^2}},
\]

and

\[
A = F_z^{\beta - 1}
\]

Be extremely careful that the space variables are \((F_0,\sigma_0)\). \((F_t,\sigma_t)\) are fixed and are related to initial conditions. This is because we used the backward equation. See [69] for more details.
This explicit expression for $p_{L_0}$ seems complicated, but the fact is that it could be obtained easily from the fact that the underlying stochastic process corresponding to $L_0$ is:

$$\begin{cases}
    dF_t = \sigma_t Af_t dW_t^F \\
    d\sigma_t = \alpha \sigma_t dW_t^\sigma \\
    dW_t^F dW_t^\sigma = \rho dt
\end{cases} \quad (5.18)$$

Such that the transition kernel for this process can be solved simply by log change of variables.

Now we use our explicit formula to compute the approximated heat kernel:

$$u(t, \cdot) = e^{rL} f \sim (e^{(ad_{L_0} + V)}Id) e^{tL_0} f \quad (5.19)$$

Here we can easily compute this express up to any order. Numerical experiments shows that the error is small enough when we truncate at:

$$u(t, \cdot) \approx (1 + tv + t^2(2ad_{L_0} + V)V/2) e^{tL_0} f \quad (5.20)$$

For the computational issue, please refer to Appendix A, where more details are provided on storing and computing the operator multiplications. The figure is a plot for the two-dimensional transition kernel for SABR model.
Figure 5.3: SABR model for $\alpha = .12, \beta = .75, \rho = -.05, F_0 = 100, \sigma_0 = .34, t = .5$
Chapter 6  
Extending the Approximating Formula to Time/Space Dependent Coefficients

6.1 Time/Space dependent coefficients

In the previous sections, we focused on the PDE with coefficients independent of time. However, in real world applications, the models always contain time-dependent coefficients, in order to embrace more flexibility to calibrate the real world data. However, this brings extra difficulty from the mathematical point of view. Unlike the ODE that one can add an auxiliary variable to make the equation coefficients time-free, since this will destroy the parabolic structure. So it is crucial to develop a method to deal with situations in which the coefficients depend on both space and time.

Specifically we consider the following problem:

\[
\begin{aligned}
\partial_t u(t, x) - L' u(t, x) &= 0 \quad \text{in } (0, \infty) \times \mathbb{R}^N \\
u(0, x) &= f(x) \quad \text{on } \{0\} \times \mathbb{R}^N,
\end{aligned}
\]  

(6.1)

for \( u, f, \) and \( g \) in suitable function spaces. In view of Duhamel’s principle, we may assume \( g = 0 \).

where

\[
L' u(x) := \sum_{i,j=1}^N a_{ij}(t, x) \partial_i \partial_j u(x) + \sum_{k=1}^N b_k(t, x) \partial_k u(x) + c(t, x) u(x),
\]  

(6.2)
Next, we generalize our approximating formula (4.12):

**Theorem 6.1.1.** The solution to 6.1 is given by:

\[
    u(t, x) = \sum_{j=0}^{\infty} P_n(t)u^0(t, x)
\]  

(6.3)

where \( P_n \) are given inductively by:

\[
    P_n(t) = \sum_{i=1}^{n} \int_{0}^{t} e^{\int_{0}^{\tau} L_0^s ds} (L_0^\tau \mathcal{P}_{n-1}(\tau)) d\tau
\]  

(6.4)

for \( n \geq 0 \), and \( P_0(t) = Id \). \( L_j^t \) are given by

\[
    L_j^t = \frac{(x - z)^j}{j!} (a^{(j)}(t, z) \partial_x^2 u + b^{(j)}(t, z) \partial_x u + c^{(j)}(t, z) u),
\]  

(6.5)

and \( u^0(t, x) \) solves the equation:

\[
    \partial_t u^0 = a(t, z) \partial_x^2 u^0 + b(t, z) \partial_x u^0 + c(t, z) u^0
\]  

(6.6)

**Proof.** The same procedure as we proved Lemma 2.2.1 and Theorem 4.2.2.

Note: as far as we can solve the problem 6.6, we do not need to ‘freeze’ \( t \). This approach is different from the approach in [37], in which the space and time are ‘frozen’ at the same time. On the other hand, if we can not solve (6.6), we can step further to freeze the time variable, but that is another problem of approximation. It makes sense to treat these approximation procedures separately.

### 6.2 A Example: Hull-White model for the Bond Price

Hull White model is a widely used short-rate model in fixed income derivatives. The instantaneous short rate of interest follows the time-dependent OU process:

\[
    dr(t) = a(t)(b(t) - r(t)) + \sigma(t)dW(t)
\]  

(6.7)

And the risk-free bond price solves the following parabolic equation:

\[
    \partial_t u(t, r) = .5\sigma^2(t)\partial_r^2 u(t, r) + a(t)(b(t) - r)\partial_r u(t, r) - ru(t, r)
\]  

(6.8)
with boundary condition:
\[ u(0, r) = 1 \]

Now using Theorem 6.1.1 we got in last section, to get:
\[
\begin{align*}
    u(t, r) &= \sum_{j=0}^{\infty} P_n(t)u^0(t, r) \\
    \text{(6.9)}
\end{align*}
\]

where \( u^0(t, x) \) solves the equation:
\[
\begin{align*}
    \partial_t u^0 &= .5\sigma(t)^2 \partial_r^2 u^0 + a(t)(b(t) - r_0)\partial_r u^0 - r_0 u^0 \\
    \text{(6.10)}
\end{align*}
\]

such that the solution is: \( u^0(t, r) = \exp(-r_0 t) \).

\( L_j^t \) are given by
\[
\begin{align*}
    L_0^t &= .5\sigma^2(t)\partial_r^2 + a(t)(b(t) - r_0)\partial_r - r_0 \\
    L_1^t &= -a(t)(r - r_0)\partial_r - (r - r_0) \\
    L_j^t &= 0, \text{ for } j \geq 2
\end{align*}
\]

Thus, we get the first few terms of \( P_n(t) \) are given by: (Take into consideration that \( u^0 \) is constant, we only need to collect those terms without polynomials and without partials)
\[
\begin{align*}
    P_1(t) &= -.5 \int_0^t \int_0^{t-\tau} a(s)(b(s) - r_0)dsd\tau \\
    \text{(6.13)}
\end{align*}
\]

The following terms are very complicated, and without certain assumptions of the forms of \( \sigma(t), a(t) \) and \( b(t) \), symbolic computational doesn’t help also. In the numerical experiment, we assume linear dependency of the parameters: \( a(t) = a_0 + a_1 t, b(t) = b_0 + b_1 t \), and \( \sigma^2(t) = \sigma_0^2 + \sigma_1^2 t \), then we have:
\[
\begin{align*}
    P_1(t) &= -.5(a_0(b_0 - r_0)t + .25(a_0b_1 + a_1(b_0 - r_0))t^2 + a_1b_1t^3/9
\end{align*}
\]

The numerical results are given below:
Figure 6.1: Hull White model with $a(t) = .2 + .01t$, $b(t) = .05 + .02t$, $\sigma^2(t) = .32^2 + .03^2t$, $r_0 = .05$, and various maturities
Chapter 7  |  
American Style Options

Most of the equity options traded on the exchanges are American style, which means that early exercise if possible. Although it is never optimal to exercise a call option earlier without dividend, there are chances that put options will be optimal to exercise earlier. Earlier exercise of options brought new difficulties from the mathematical point of view. Even for the Black-Scholes model, there are no closed form solutions for American options. Numerical methods seems unavoidable. In numerical methods, one always needs to divided the time horizon to small intervals, and decide if early exercise is optimal or not in each interval, in a dynamic programming fashion. This procedure perfectly falls into our treatment, since our short-time approximation does not add more technical difficulty-we get small time for free! We use the CEV model as an example to show how we treat the problem.

7.1 American put option for the CEV model with zero interest rate

In order to approximate the American option, we need to first consider the heat kernel expansion of the CEV model:

\[ \partial_t u = .5\sigma^2 \exp(2(\alpha - 1)x)(\partial_x^2 - \partial_x) \]  

(7.1)

We follow the treatment we did in Chapter 3:

\[ u(t, x) = \sum_{j=0}^{\infty} P_n(2z t) u^0(t, x) \]  

(7.2)
where $u^0$ solves the log Black-Scholes formula:

$$
\partial_t u^0(t, x) = .5\sigma^2 \exp(2(\alpha - 1)z)(\partial_x^2 - \partial_x)u^0(t, x)
$$

such that

$$
 u^0(t, x) = \text{Call}_{BS}(S = \exp(x), t, \sigma S_0^{\alpha - 1}, K, r = 0),
$$

For this case, we get very clean expression of $L_j$:

$$
 L_j = \frac{s^j(x - z)^j}{2j!} (\partial_x^2 - \partial_x)
$$

where $s = 2(\alpha - 1)$.

While $P_j(t)$ can be computed without symbolic computation. We emphasize that we do not fix $z = \log(S_0)$, but setting $z = .5(\log(S_0 + K)$, and numerical experiments shows that this gives better approximating results.

From Figure 7.1a, we see that the 10th order approximation maintains an error within .1%, even for deep out of the monty puts.

(a) absolute price error  
(b) percentage error

Figure 7.1: CEV model European put option price, with $t = 1.5, \sigma = .34, \alpha = .65, S_0 = 100, r = 0$, and various strikes.
7.2 Heat kernel expansion for CEV-first attempt

It is straightforward to expand the heat kernel of the CEV model:

\[ G_{CEV}(t; x, y) = \sum_{j=0}^{\infty} P_n(2a(z)t)G^0(t; x, y) \] (7.5)

where \( G^0 \) solves the heat equation:

\[ \partial_t G^0(t; x, y) = .5\sigma^2 \exp(2(\alpha - 1)z)(\partial^2_x - \partial_x)G^0(t; x, y) \] (7.6)

with initial condition

\[ G^0(0; x, y) = \delta_y(x) \]

where \( \delta \) is the Dirac function. Letting \( V_z = \sigma^2 \exp(2(\alpha - 1)z) \), we get:

\[ G^0(t; x, y) = \exp\left(-\frac{(x-.5V_zt-y)^2}{2V_zt}\right) \]

Note that this is the log version of the heat kernel. The heat kernel for the original is:

\[ G_{originalCEV}(t; S, S_t) = G_{logCEV}(t; \log(S_0), \log(S_t))/S_t \]

We present the numerical experiment for the heat kernel approximation in Figure 7.2.

Next, for American style option, we divide the time range into several small intervals. In each interval, we price options by deciding if the option is exercised immediately, or wait to next time step. Mathematically, we let \( t_j = jt/M \), for \( j = 1, \cdots, M \), for a large integer \( M \), and for each interval \( [t_{j-1}, t_j] \), we solve the problem of the following:

\[ u(t_{j-1}, x) = \max\{\exp(x) - K, \int_{-\infty}^{\infty} \rho_{logCEV}(t_j - t_{j-1}; x, k)u(t_j - t, k)dk\} \] (7.7)

which will cause discretization error. But this error is not avoidable even for Black-Scholes.

The numerical experiment is:

From the figure, we notice that second order approximation of the heat kernel is necessary. We omit here some details of the numerical integrations. In later chapter, we will get back to the problem of approximating the heat kernel.
Figure 7.2: CEV model for $\alpha = .75, S_0 = 100, \sigma = .34, t = 1.5, r = 0$
short time expansion for American put option of CEV model
forward put option price
0th order
2nd order
4th order
closed form solution for European put

Figure 7.3: CEV model for $\alpha = .65, Strike = 100, \sigma = .34, t = 2, r = .0$, with forward ranging from 60 to 140
Starting with Merton’s seminal paper [14] and up to the present date, various aspects of models with jumps have been studied in the academic finance community. In the last decade, also the research departments of major banks started to accept jump-diffusions and Lévy processes as a valuable tool in their day-to-day modeling. This increasing interest to jump models in finance is mainly due to the following reasons.

First, in a model with continuous paths like a diffusion model, the price process behaves locally like a Brownian motion and the probability that the stock moves by a large amount over a short period of time is very small, unless one fixes an unrealistically high value of volatility. Therefore, in such models the prices of short term out of the money (OTM) options should be much lower than what one observes in real markets. On the other hand, if stock prices are allowed to jump, even when time to maturity is very short, there is a non-negligible probability that after a sudden change in the stock price the option will move in the money. Second, from the point of view of hedging, continuous models of stock price behavior generally lead to a complete market or, at least, to a market, that can be made complete by adding one or two additional instruments, like in stochastic volatility models [12]. Since in such a market every terminal payoff can be exactly replicated, options are redundant assets, and the very existence of traded options becomes a puzzle. The mystery is easily solved by allowing for discontinuities: in real markets, due to the presence of jumps in the prices, perfect hedging is impossible and options enable the market participants to hedge risks that cannot be hedged using the underlying only. From a risk management perspective, jumps allow to quantify and take into account the risk of strong stock price movements over short time intervals, which appears non-existent in the diffusion framework. The last and probably the strongest argument for using discontinuous models is simply the presence of jumps in observed prices.

In order to understand the jump models, we need to first introduce the Lévy Process.
8.1 Review of Lévy processes

We now review a few shall needed concepts and definitions following mostly [70]. Let us recall that a stochastic process \( \{X_t : t \geq 0\} \) on \( \mathbb{R}^N \) a probability space \((\Omega, \mathbb{P})\) is a Lévy process if the following conditions are satisfied.

(i) For any choice of \( n \geq 1 \) and \( 0 \leq t_0 < t_1 < \cdots < t_n \), the random variables \( X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}} \) are independent. (This property is called the independent increments property).

(ii) \( X_0 = 0 \) almost surely (a.s.).

(iii) The distribution of \( X_{s+t} - X_s \) does not depend on \( s \) (temporal homogeneity or stationary increments property).

(iv) It is stochastically continuous, that is, for every \( t \geq 0 \) and \( \epsilon > 0 \), we have that

\[
\lim_{s \to t} \mathbb{P}[|X_s - X_t| > \epsilon] = 0
\]

(v) There is \( \Omega_0 \) with \( \mathbb{P}[\Omega_0] = 1 \) such that, for every \( \omega \in \Omega_0 \), \( X_t(\omega) \) is right-continuous in \( t \) for \( t \geq 0 \) and has left limits in \( t \) for \( t > 0 \). (A process with this property is called cádlág.)

We include now some typical examples of Lévy processes:

8.1.1 The Poisson Process

Take a sequence \( \{\tau_i\}_{i \geq 1} \) of independent exponential random variables with parameter \( \lambda \). (Recall that a random variable \( \tau \) is exponential with parameter \( \lambda \) if, by definition, its cumulative distribution function is \( \mathbb{P}[\tau \geq y] = e^{-\lambda y} \).) Let \( T_n = \sum_{i=1}^{n} \tau_i \). Then the process

\[
N_t = \sum_{n \geq 1} \mathbb{1}_{t \geq T_n}
\]

is called the Poisson Process with parameter \( \lambda \). At every time \( t > 0 \), \( N_t \) has the Poisson distribution with parameter \( \lambda t \), that is, it is integer-valued and

\[
\mathbb{P}[N_t = n] = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.
\]
The characteristic function of the poisson process \( X \) with parameter \( \lambda \) is
\[
\phi_X(u) \equiv \mathbb{E}[e^{iuX}] = \exp \left\{ \lambda t(e^{iu} - 1) \right\}.
\]

### 8.1.2 Compound Poisson processes

For financial applications, it is of little interest to have a process with a single possible jump size. The compound Poisson process is a generalization where the waiting times between jumps are exponential but the jump sizes can have an arbitrary distribution. More precisely, let \( N \) be a Poisson process with parameter \( \lambda \) and \( \{Y_i\}_{i \geq 1} \) be a sequence of independent random variables with law \( f \). The process
\[
X_t = \sum_{i=0}^{N_t} Y_i
\]
is called a compound Poisson process. The characteristic function is known and has the form
\[
\mathbb{E}[e^{iuX_t}] = \exp \left\{ t \lambda \int_{\mathbb{R}} (e^{iux} - 1)f(dx) \right\}.
\]

### 8.1.3 Jump-diffusion and Lévy processes

Combining a Brownian motion with drift and a compound Poisson process, we obtain the simplest case of a jump-diffusion-a process which sometimes jumps and has a continuous but random evolution between the jump times:
\[
X_t = \mu t + \sigma B_t + \sum_{i=0}^{N_t} Y_i.
\] (8.1)

The best known model of this type in finance is the Merton model, where the stock price is \( S_t = S_0 e^{X_t} \) with \( X_t \) as above and the jumps \( \{Y_i\} \) have Gaussian distribution \( f \). The characteristic function of the process (8.1) is
\[
\mathbb{E}[e^{iuX_t}] = \exp \left\{ t \left( i\mu u - \frac{\sigma^2 u^2}{2} + \lambda \int_{\mathbb{R}} (e^{iux} - 1)f(dx) \right) \right\}.
\] (8.2)

Actually, the following theorem holds [70]
Theorem 8.1.1. If \( X_t \) is a Lévy process on \( \mathbb{R} \), then

\[
\mathbb{E}[e^{iuX_t}] = \exp\left\{ t \left( i\mu u - \frac{\sigma^2 u^2}{2} \right) + \lambda \int_{\mathbb{R}} (e^{iux} - 1 - \mathbb{1}_{\{|x|\leq 1\}} iux ) \nu(dx) \right\}, \tag{8.3}
\]

where \( \nu \) is a measure on \( \mathbb{R} \) satisfying

\[
\nu(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}} \min(1, x^2) \nu(dx) < \infty. \tag{8.4}
\]

The measure \( \nu \) is called the Lévy measure of \( X_t \). Now we can investigate more complicated Lévy processes that are not intuitive as the compound Poisson processes, which could be described by their characteristic functions.

8.1.4 Gamma processes

The Gamma process with parameters \( \lambda \) and \( c \) is a Lévy process with the law \( p_t \) or \( X_t \) is the gamma law with parameters \( \lambda \) and \( ct \):

\[
p_t(x) = \frac{\lambda^c t}{\Gamma(ct)} x^{ct-1} e^{-\lambda x}. \tag{8.5}
\]

where \( \Gamma(x) = \int_{\mathbb{R}^+} e^{-tx} dt \). The Gamma process is an increasing Lévy process (also called subordinator). Its characteristic function has a very simple form:

\[
\mathbb{E}[e^{iuX_t}] = \frac{1}{(1 - iu/\lambda)^{-ct}}. \tag{8.6}
\]

The Gamma process is the building block for a very popular jump model, the “variance gamma process” discussed next.

8.1.5 Variance Gamma process

The Variance Gamma process is constructed by taking a Brownian motion with drift and changing its time scale with a Gamma process:

\[
Y_t = \mu X_t + \sigma B_{X_t}. \tag{8.7}
\]
where \( X_t \) is a standard Gamma process with parameters \( \lambda \) and \( c \). The characteristic function of the Variance Gamma process is

\[
\mathbb{E}[e^{iuY_t}] = \left( 1 + \frac{\sigma^2 u^2}{2} - i\mu c u \right)^{-ct}.
\] (8.8)

### 8.1.6 Inverse Gaussian Process

The Inverse Gaussian Process with parameters \( \lambda \) and \( c \) is a Levy process with the law \( p_t \), or \( X_t \) is the inverse Gaussian law with parameters \( \lambda \) and \( ct \):

\[
p_t(x) = \frac{ct}{x^{3/2}} e^{-\frac{(\sqrt{\lambda}x - \sqrt{\pi}ct)^2}{x}}
\] (8.9)

The inverse Gaussian process is an increasing Levy process, and it is not hard to verify that this process is closely related to the first arriving time of a Brownian motion with drift. Its characteristic function has the form:

\[
\mathbb{E}[e^{iuX_t}] = \exp(-2ct\sqrt{\pi}(\sqrt{\lambda} - iu - \sqrt{\lambda}))
\] (8.10)

### 8.1.7 Normal Inverse Gaussian process

The Normal Inverse Gaussian process is constructed by taking a Brownian with drift and changing its time scale with an Inverse Gaussian process:

\[
Y_t = \theta X_t + \sigma B_t
\] (8.11)

The characteristic function of the Normal Inverse Gaussian process is:

\[
\mathbb{E}(e^{iuY_t}) = \exp\left(t\frac{1 - \sqrt{1 + u^2\sigma^2\kappa - 2i\theta \mu \kappa}}{\kappa}\right)
\] (8.12)

### 8.1.8 Strictly \( \alpha \)-stable processes

If we take

\[
\nu(dx) = \frac{A}{|x|^{\alpha+1}}
\]

in (8.1.1), where \( A \) is a positive constant, \( 0 \leq \alpha \leq 2 \), then we get the characteristic function of the so called strictly \( \alpha \)-stable process \( X_t \). One can easily check that this measure satisfies
(8.4). It is stable because it is the ‘attractor’ for properly normed sums of independent and identically-distributed random variables. Please refer to Chapter of [70] for further details. The characteristic function of strictly $\alpha$-stable process is

$$\mathbb{E}[e^{iuX_t}] = e^{-ct|u|^\alpha}$$ (8.13)

The consequence of stable process is that any change of time scale for the process has the same effect as some change of spatial scale. This property is called self-similarity of a stochastic process.

For strictly $\alpha$-stable process $X_t$, we have

$$X_{\epsilon t} \overset{d}{=} cX_t$$ (8.14)

where $\overset{d}{=} \text{ means equal in the sense of distribution.}$

### 8.2 Exponential Lévy models

To ensure positivity as well as the independence and stationarity of log-returns, stock prices are usually modeled as exponentials of Lévy processes:

$$S_t = S_0 e^{X_t}.$$ (8.15)

where $X_t$ is a Lévy process.

#### 8.2.1 A survey of popular jump models

1) General jump diffusion model:

$$dS_t = \mu S_t dt + \sigma S_t dW_t + S_- dJ_t$$ (8.16)

where $W_t$ is a Brownian motion and

$$J_t = \sum_{i=1}^{N_t} Y_i$$ (8.17)

where $N_t$ is Poisson process with jump intensity $\lambda$. $\mu$ is chosen such that $e^{-rt} S_t$ is a martingale. $Y_i$ are iid random variables. According to the distribution of $Y_i$, there are several models:
1.1) Poisson jump model:

\[ Y_i = \begin{cases} 
  d_+ & \text{with probability } p \\
  d_- & \text{with probability } q = 1 - p
\end{cases} \quad (8.18) \]

1.2) Merton’s jump model [14]: \( Y_i \sim N(\alpha, \delta) \)

1.3) Kou’s model [15] \( \log(Y_i) \) has an asymmetric double exponential distribution with the density

\[ f(y) = p\eta_1 e^{-\eta_1 y} 1_{y \geq 0} + q\eta_2 e^{\eta_2 y} 1_{y < 0} \quad (8.19) \]

where \( p, q \geq 0, p + 1 = 1, \) and \( \eta_1 > 1, \eta_2 > 0. \)

2) Variance Gamma(VG) model

In [?], the authors introduced the Variance Gamma model, which is a Brownian motion subordinated by a Gamma process. It is proved that this process can be written as the difference of two Levy processes with tempered stable jump measures.

Let \( G(t, \nu) \) be an independent gamma process with mean rate unity and variance rate \( \nu. \) The variance gamma process is defined by:

\[ X_{VG}(t; \sigma, \nu, \theta) = \theta G(t; \nu) + \sigma W_{G(t; \nu)} \quad (8.20) \]

And the price process is defined as:

\[ S_t = S_0 \exp(\mu t + X_{VG}(t; \sigma, \nu, \theta)) \quad (8.21) \]

3) CGMY model On the base of VG, [16] proposed a new model:

\[ S_t = S_0 \exp(\mu t + \sigma W_t + X_t) \quad (8.22) \]

where \( X_t \) is a Levy process without drift and diffusion, and the jump measure is given by:

\[ k_{CGMY}(x) = \begin{cases} 
  C \exp\left(-\frac{G|x|}{|x|+\gamma}\right) & \text{for } x < 0 \\
  C \exp\left(-\frac{Mx}{|x|+\gamma}\right) & \text{for } x > 0
\end{cases} \quad (8.23) \]

4) Normal Inverse Gaussian (NIG) model

Another important model is the Normal Inverse Gaussian (NIG) model introduced in [71], in which a Brownian motion is subordinated by an Inverse Gaussian process. The Inverse Gaussian process is the stopping time process of a Brownian motion with positive drift.
Please refer to Chapter 4 of [72] for more details.

\[ S_t = S_0 \exp(\mu t + \sigma W_t + X_t) \]  
(8.24)

where \( X_t \) is a normal inverse Gaussian process with parameters \( \nu, \alpha, \beta, \) and \( \delta. \)

5) Bates volatility model

In [73], Bates proposed a stochastic volatility model, with jumps in stock price process:

\[
\begin{aligned}
    dS_t &= (r - \delta)S_t - dt + \sqrt{V_t}S_t - dW_t + S_t - dJ_t \\
    dV_t &= \kappa(\theta - V_t) + \sigma_v \sqrt{V_t}dZ_t 
\end{aligned}
\]

where \( W_t \) and \( Z_t \) are Brownian motions such that \( dW_t dZ_t = \rho dt, \) where \(-1 < \rho < 1\). \( J_t \) is the compound Poisson process with intensity \( \lambda \) and independent jumps \( Y \) with

\[
\log(1 + Y) = N(\log(1 + \alpha) - \frac{1}{2} \beta^2, \beta^2)
\]
(8.25)

6) General Affine models

In [61], the authors proposed the general approach to affine models:

\[ S_t = S_0 e^{X_1} \]  
(8.26)

and the SDE satisfied by \( X_t \) is given by:

\[
dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + dZ_t
\]
(8.27)

where \( dZ_t \) is a simple jump process(we will explain why this is simple later). with the following affine assumption:

\[
\mu(x) = K_0 + K_1 x, \text{ for } K = (K_0, K_1) \in \mathbb{R}^N \times \mathbb{R}^{N \times N} \\
(\sigma(x)\sigma(x)^T)_{i,j} = (H_0)_{i,j} + (H_1)_{i,j} \cdot x, \text{ for } H = (H_0, H_1) \in \mathbb{R}^{N \times N} \times \mathbb{R}^{N \times N \times N} \\\n\lambda(x) = l_0 + l_1 \cdot x, \text{ for } l = (l_0, l_1) \in \mathbb{R} \times \mathbb{R}^N
\]
8.3 Outside Fourier Transform regime: commutator method applied to Kou’s model with local volatility

In this section, we treat a generalized version of Kou’s model in the following form:

\[ dY(t) = \sigma dW(t) + (r - 0.5\sigma^2 - \lambda h(Y(t)))dt + h(Y(t))dJ(t) \] (8.28)

where \( h(y) \) is the local volatility of the jump, and \( J(t) \) is the jump in Kou’s model with intensity \( \lambda \) and jump measure given by:

\[ f_{Kou}(y) = p\eta_1 e^{-\eta_1 y}1_{y \geq 0} + q\eta_2 e^{\eta_2 y}1_{y < 0} \] (8.29)

where \( p, q \geq 0, p + 1 = 1, \) and \( \eta_1 > 1, \eta_2 > 0. \) \( m = \frac{p}{\eta_1 - 1} - \frac{q}{\eta_2 + 1}, \) which is obtained by Levy – Kihintchine formula, to compensate the jump process, such that the discounted stock process is a martingale under risk neutral measure.

The corresponding partial differential-integral equation reads:

\[ \partial_t u(t, x) = 0.5\sigma^2 \partial_x^2 u(t, x) + (r - 0.5\sigma^2 - \lambda h(x))\partial_x u(t, x) + \lambda h(x) \int_{-\infty}^{\infty} (u(t, x+y) - u(x)) f_{Kou}(y) dy \] (8.30)

We follow the procedure of setting

\[ L_0 = 0.5\sigma^2 \partial_x^2 + (r - 0.5\sigma^2)\partial_x \]
\[ L_j = \lambda \frac{h^{(j)}(z)(x-z)^j}{j!} \mathcal{I} \text{ for } j \geq 1 \]

where

\[ \mathcal{I} = \int_{-\infty}^{\infty} (u(t, x+y) - u(x)) f_{Kou}(y) dy - m\partial_x \]

And the first several few terms of \( \mathcal{P}_n \) are:

\[ \mathcal{P}_1(t) = \lambda h'(z)(t(x-z) + 0.5t^2(\sigma^2 \partial_x + (r - 0.5\sigma^2)))\mathcal{I} \]
\[ \mathcal{P}_2(t) = \ldots \text{complicated terms, but matlab code can easily compute them.} \]

The solution is given by:

80
\[ u(t, x) = \sum_{j=0}^{\infty} \mathcal{P}_n(t) u^0(t, x) \]  

(8.31)

where \( u^0 \) solves:

\[ \partial_t u^0 = L_0 u^0 \]  

(8.32)

with initial condition for certain payoff structure, which is closed related to the Black-Scholes solution. Also notice that we have integral terms in \( \mathcal{P}_n \), such that the solution can be viewed as a weighted average of Black-Scholes options for a range of stock prices. Comparing this with the approximation formula we got from Chapter 3 for pure diffusion process (where the Feynman-Kac formulae gives us a partial differential equation, and where the solution depends only on the Black-Scholes at one point plus all the mathematical derivatives at that point), we understand in this sense why these models are called non-local. And the numerical experiment results are in the following:

Figure 8.2a, we see that when there are nontrivial local volatility, there is no longer Fourier transform solutions, and we can only compare the expansive solution to the Monte Carlo solution.

8.4 Local volatility jump-diffusion models with stochastic volatility

In this section, we illustrate the general approach for local volatility version of jump-diffusion models for exponential Levy process. Specifically, we show how to approximate the local volatility version of Bate’s model:

\[ \begin{align*}
    dS_t &= (r - \delta)S_t dt + \sqrt{V_t} S_t dW_t + S_t \gamma dJ_t \\
    dV_t &= \kappa(\theta - V_t) + \sigma \sqrt{V_t} dZ_t
\end{align*} \]

where \( W_t \) and \( Z_t \) are Brownian motions such that \( dW_t dZ_t = \rho dt \), where \(-1 < \rho < 1\). \( J_t \) is the compound Poisson process with intensity \( \lambda \) and independent jumps \( Y \) with

\[ \log(1 + Y) = N(\log(1 + \mu_J) - \frac{1}{2} \sigma^2_J, \sigma^2_J) \]  

(8.33)

0 < \gamma < 1, to account for the leverage effect from the jump process: when the stock price goes up, the jump becomes fewer.
Letting $x = \log(S)$ and $v(t, x, V) = u(t, e^x, V)$, we have the following partial-integral differential equation reads:

$$
\partial_t u = .5V\partial_x^2 u(t, x, V) + .5\sigma_V^2 V\partial^2_V u + \rho \sigma_V V\partial_{xV} + (r - .5V - \lambda m \exp((\gamma - 1)y))\partial_x u \\
+ \kappa(\theta - V)\partial_V u - ru - \lambda \exp((\gamma - 1)y)\mathcal{I} u
$$

where $\mathcal{I} u = \int_{-\infty}^{\infty} (u(t, x + y, V) - u(t, x, y)) f_{Bates}(y) dy$, and $f_{Bates}(y)$ is the jump measure.
(a) absolute price error  

(b) percentage error  

Figure 8.2: Kou’s model for $\lambda = .1, S_0 = 100, \sigma = .34, p = .4, \eta_1 = 2.1, \eta_2 = 1.9, t = .75, r = .05$, and local volatility $h(x) = \exp(-.25(x - \log(S_0)))$, compared to Monte Carlo simulation solution for various strikes of the log stock process, which is given by:

$$f_{Bates}(y) = \frac{\exp(-\frac{(y-\mu_j)^2}{2\sigma_j^2})}{\sqrt{2\pi\sigma_j^2}}$$

and $m = \exp(\mu_j + .5\sigma_j^2) - 1$, is the drift to compensate for the jump, such that the discounted stock process is a martingale under risk neutral measure.

We have:

$$L_0 = .5V_0\partial_x^2 + .5\sigma_V^2V_0\partial_V^2 + \rho\sigma_VV_0\partial_{xV} + (r - .5V_0)\partial_x + \kappa(\theta - V_0)\partial_V - ru$$

$$L_1 = .5(V - V_0)\partial_x^2 + .5\sigma_V^2(V - V_0)\partial_V^2 + \rho\sigma_V(V - V_0)\partial_{xV}$$

$$- (.5(V - V_0) + \lambda m\epsilon(\gamma - 1)y_0)\partial_x - (\kappa(V - V_0))\partial_V + \lambda\epsilon(\gamma - 1)y_0\mathcal{I}u$$

$$L_j = \frac{((\gamma - 1)(y - y_0))^{j-1}}{(j - 1)!}(-\lambda m\epsilon(\gamma - 1)y_0\partial_x + \lambda\epsilon(\gamma - 1)y_0\mathcal{I}u)$$
(a) European call option price  

(b) percentage error

Figure 8.3: Bates model European call price, with $\kappa = 4, \theta = .24, \sigma_V = .23, \rho = -.03, \lambda = .05, \mu_J = .0001, \sigma_J = .01, t = .25, V_0 = .34^2, S_0 = 30, r = .05$ and various strikes.

The numerical computation is straight forward, we present the results in Figure 8.3a with original Bates model, and 8.4a with local volatility version of Bates model:
Figure 8.4: Bates model European call price, with $\kappa = 4, \theta = .24, \sigma_V = .23, \rho = -.03, \lambda = .1, \mu_J = .0001, \sigma_J = .1, t = .25, V_0 = .34^2, S_0 = 30, r = .05,$, with local volatility in the form of $h(x) = \exp((.7 - 1)(x - \log(S_0)))$, and various strikes.
Chapter 9
Approximate the Implied Volatilities/Heat Kernels for Various Models

People like Black-Scholes, and always want to connect the option price with Black-Scholes models. Since there is a one-to-one correspondence between option price and the volatility in Black-Scholes model, people are interested in finding the corresponding volatility such that the Black-Scholes gives the option price. This implied volatility, also known as ‘the wrong number when substituted into the wrong formula, gives the right answer’, is of most interest by practitioners. Actually, in industry practice, quants are always trying to calibrate the implied volatility surface, rather than the option price spaces. In this section, we will derive the implied volatility for several popular models. At the same time, we are also interested in the approximation of transition kernel, and characteristic functions.

9.1 Approximation to the implied volatility/heat kernel of CEV model

we consider again the CEV model: We consider the CEV model, after the log change of the state variable, we get:

\[ \partial_t u(t, x) = 0.5\sigma^2 \exp(2(\alpha - 1)x)(\partial_x^2 - \partial_x)u \]  
(9.1)
we want to solve \( u \) at \((t, z = \log(S_0))\), now we change use (4.18) we got in last section, to get:

\[
u(t, x) = \sum_{j=0}^{\infty} P_n(a(z_0)t)u^0(t, x) \tag{9.2}\]

where \( a(z) = \sigma^2 \exp(2(\alpha - 1)z_0) \), and \( u^0 \) solves the log Black-Scholes formula:

\[
\partial_t u^0(t, x) = 0.5\sigma^2 \exp(2(\alpha - 1)z)(\partial^2_x - \partial_x)u^0(t, x) \tag{9.3}
\]

such that

\[
u^0(t, x) = \text{Call}_{BS}(S = \exp(x), t, \sigma e^{(\alpha-1)z_0}, K, r = 0),
\]

For this case, we get very clean expression of \( L_j \):

\[
L_j = \frac{s^j(x - z)^j}{2^j!} (\partial^2_x - \partial_x) \tag{9.4}
\]

where \( s = 2(\alpha - 1) \). and \( P_j(t) \) can be computed explicitly, because of the simplicity of expressions of \( L_j \). Without the terms of \((x - z_0)\), we have:

\[
P_1(t) = st^2(2\partial^3_x - 3\partial^2_x + \partial_x)/8,P_2(t) = s^2(t^2(\partial^2_x - \partial_x)/8 + t^3(5\partial^4_x - 10\partial^3_x + 6\partial^2_x - \partial_x))/24 + t^4(4\partial^6_x - 12\partial^5_x + 13\partial^4_x - 6\partial^3_x + \partial^2_x)/128)
\]

It seems that it is straightforward to invert the call option price formula to get an expansive solution to the implied volatility. However, this approach is not efficient, from the point of view that taking derivatives with respect to \( \sigma \) will give complicated terms; and this is not beautiful, from the point of view of mathematical harmony. We take the following approach: Assume that for strike \( K \) and time \( t \), the implied volatility has the expansive expression:

\[
\sigma^2_{Imp}(S_0, t, K) = \tilde{\sigma}^2(1 + \sum_{j=1}^{\infty} s^j q_j) \tag{9.5}
\]

note that \( s = 2(\alpha - 1) \), this is the connection to the CEV model. Then we substitute this into Black-Scholes formula:

\[
\partial_t u = 0.5\sigma^2_{Imp}(\partial^2_x - \partial_x)u = 0.5\tilde{\sigma}^2(1 + \sum_{j=1}^{\infty} s^j q_j)(\partial^2_x - \partial_x)u \tag{9.6}
\]
Now we recall the small parameter approximation formula (4.18) to get:

\[ u(t, x) = \sum_{j=0}^{\infty} \tilde{P}_n(\tilde{a}(z)t)u^0(t, x) \]  

(9.7)

where \( \tilde{a}(z) = \tilde{\sigma}^2 \), and \( u^0 \) solves the log Black-Scholes formula:

\[ \partial_t u^0(t, x) = .5\tilde{\sigma}^2(\partial_x^2 - \partial_x)u^0(t, x) \]  

(9.8)

such that

\[ u^0(t, x) = \text{Call}_{BS}(S = \exp(x), t, \tilde{\sigma}, K, r = 0), \]

For this case, we get very clean expression of \( L_j \):

\[ \tilde{L}_j = \frac{s^j}{2}(\partial_x^2 - \partial_x) \]  

(9.9)

and \( \tilde{P}_j(t) \) can be computed explicitly:

\[ \tilde{P}_1(t) = \frac{s^1}{2}(\partial_x^2 - \partial_x) \]
\[ \tilde{P}_2(t) = \frac{s^2}{2}(\partial_x^2 - \partial_x) + \frac{s^2q_1^2}{8}(\partial_x^2 - \partial_x)^2 \]

and so on. Now we have two formal expansions (9.2) and (9.7), since they leads to the same solution at \((t, x_0 = \log(S_0))\), we can compare them to get the expansion for the implied vol as follows: By comparing the first three terms of the expansion, we get:

\[ \tilde{\sigma} = \sigma \exp((\alpha - 1)z) \]  

(9.10)

\[ q_1 = (x_0 - z) - .5d_1\sqrt{T} + .25T = (x_0 - z) - .5k \]  

(9.11)

\[ q_2 = .5(x_0 - z)^2 + (x_0 - z)(-.5d_1\sqrt{T} + .25T) + \frac{T^2 - 5d_1\sqrt{T}^3 + (1 + 5d_1^2)T}{48} \]

\[ = .5(x_0 - z)^2 - .5(x_0 - z)k + \frac{-2.25T^2 + 5k^2 + T}{48} \]  

(9.12)

(9.13)

where \( T = \sigma^2 \exp(2(\alpha - 1)z)t \), and \( d_1 = \frac{k + .5T}{\sqrt{T}} \), \( k = \log S_0 - \log K \) is the log-moneyness.

We have different choice of \( z \). In [23], \( z = \log(S_0 + \frac{K}{2}) \), there are other possible choices:
Figure 9.1: CEV model European call price, with $t = 5.25$, $\sigma = .34$, $\alpha = .5$, $S_0 = 90$, $r = 0$, and various strikes

$z = \frac{\log(S_0) + \log(K)}{2}$ as the geometric mean, which makes $q_1 = 0$, and $z = \log(F_0)$ as the current forward price. We see from Figure 9.1, that 2nd order expansion with geometric mean is the best choice. We also include the approximation formula in [23] and found that my expansion is arguably better than theirs.

We given explicitly the geometric mean implied volatility approximation for the CEV model, we use the notations in Hagan and Woodward(1999) to show my respect to their original ideas:

\begin{equation}
    dF(t) = aF^{\beta}dW(t)
\end{equation}

we obtain:

\begin{equation}
    \sigma_B^2 = \frac{a^2}{f_{av}^{2-2\beta}} \left( 1 + (\beta - 1)^{2}\left( \frac{T - k^2}{12} - \frac{T^2}{48} \right) + O((\beta - 1)^4) + \cdots \right)
\end{equation}

or

\begin{equation}
    \sigma_B = \frac{a}{f_{av}^{-\beta}} \left( 1 + (\beta - 1)^{2}\left( \frac{T - k^2}{24} - \frac{T^2}{96} \right) + O((\beta - 1)^4) + \cdots \right)
\end{equation}
where $f_{av} = \sqrt{F_0 K}$, $k = \log(F_0/K)$ is the log-moneyness, and $T = \frac{a^2 t}{f_{av}^{2-\beta}}$ is the scaled time to maturity.

Notice that exactly the same method could be used to find the implied volatility for the heat kernel expansion, which do not have to be the same for the implied volatility for the option pricing. We have the following results:

**Theorem 9.1.1.** The heat kernel for the CEV model will be approximated by:

$$
G_{CEV}(t; F_0, F_t) = \exp(\sigma_z^2 t A) \frac{\exp\left(-\frac{(\log F_0 - \log F_t - 5\sigma_z^2 t)^2}{2\sigma_z^2 t}\right)}{\sqrt{2\pi t \sigma_z F_t}}
$$

where $\sigma_z = \frac{a}{f_{av}^{1-\beta}}$ and $f_{av} = \sqrt{F_0 K}$. $A$ has the following expansive expression:

$$
A = a_1(\beta - 1) + a_2(\beta - 1)^2 + a_3(\beta - 1)^3 + a_4(\beta - 1)^4 + \cdots
$$

where

$$
a_1 = \frac{k}{T}
$$

$$
a_2 = \frac{T^2}{384} - \frac{T}{16} - \frac{k^4}{24T^2} + \frac{1}{8}
$$

$$
a_3 = 0
$$

$$
a_4 = \frac{-k^6}{720} - \frac{T(3T^3 - 120T^2 + 1040T - 960)}{15360}
$$

where $T = \frac{a^2 t}{f_{av}^{2-\beta}}$ is the rescaled time to maturity, $k = \log F_0 - \log K$ is the log-moneyness.

We find that there are only volatility smirks, instead of volatility smiles. We need to explore more complicated models to recover the volatility smiles observed on the market.

### 9.2 Approximation to the implied volatility/heat kernel of SABR model

We are now ready to deal with the SABR model, since we know how to deal with higher dimensional case, and since we already have the result from CEV, which is the case when $\alpha = 0$ in SABR model. Generally speaking, there are nothing conceptional new in this section,
however, the result would be very important. We will compare our series expansion result in [24]

Recall the SABR model (cite the paper by Hagan):

\[
\begin{align*}
    dF_t &= \sigma_t F_t^\beta dW_t \\
    d\sigma_t &= \alpha \sigma_t dZ_t
\end{align*}
\]

Here \(F_t\) is the forward price, and \(dW_t dZ_t = \rho dt\).

Now we write down the Feynman-Kac after the log change of forward:

\[
\begin{align*}
    \partial_t u &= .5 \sigma^2 \exp(2(\beta - 1)x)(\partial_x^2 - \partial_x)u + \rho \alpha \sigma^2 \exp((\beta - 1)x)\partial_x u + .5 \alpha^2 \sigma^2 \partial_x^2 u \\
    \partial_t u^0(t; x, \sigma) &= .5 \sigma_z^2 \exp(2(\beta - 1)x_z)(\partial_{x_z}^2 - \partial_{x_z})u^0(t; x, \sigma)
\end{align*}
\]

We have

\[
u(t; x, \sigma) = \sum_{n=0}^{\infty} P_n(t)u^0(t; x, \sigma)
\]

where \(u^0(t; x, \sigma)\) solves the Black formula:

\[
\begin{align*}
    \partial_t u^0(t; x, \sigma) &= .5 \sigma_z^2 \exp(2(\beta - 1)x_z)(\partial_{x_z}^2 - \partial_{x_z})u^0(t; x, \sigma)
\end{align*}
\]

such that the volatility is \(V_z = \sigma_z \exp((\beta - 1)x_z)\) for the Balck formula, where \(x_z\) and \(\sigma_z\) are freezes state variable which will be specified later.

\(P_n(t)\) are inductively given by Lemma 2.2.1, or recursively given by (2.15), where \(L_j\) are given by:

\[
\begin{align*}
    \frac{L_0}{\nu_z} &= .5(\partial_x^2 - \partial_x) + \frac{\rho \alpha \sigma_z^2}{\nu_z} \partial_x \sigma + .5 \alpha^2 \frac{\sigma_z^2}{\nu_z} \partial_x^2 \\
    \frac{L_1}{\nu_z} &= (\sigma - \sigma_z) + s(x - x_z)(\partial_x^2 - \partial_x) + \frac{\rho \alpha}{\nu_z} (2(\sigma - \sigma_z) + s\sigma_z(x - x_z)) \partial_x \sigma + \frac{\alpha^2 \sigma_z}{\nu_z} \partial_x^2 \\
    \frac{L_2}{\nu_z} &= \frac{(\sigma - \sigma_z)^2}{2 \sigma_z^2} + \frac{2s(\sigma - \sigma_z)(x - x_z)}{\sigma_z} + s^2(x - x_z)^2(\partial_x^2 - \partial_x) \\
    &\quad + \frac{\rho \alpha}{\nu_z} (s^2 \sigma_z(x - x_z)^2 \frac{2}{2} + 2s(x - x_z)(\sigma - \sigma_z) + \frac{(\sigma - \sigma_z)^2}{\sigma_z}) \partial_x \sigma + \frac{\alpha^2 (\sigma - \sigma_z)^2}{2 \nu_z} \partial_x^2 \\
    \frac{L_j}{\nu_z} &= \frac{(2s)^j(x - x_z)^j}{2j!} + \frac{(2s)^{j-1}(x - x_z)^{j-1}(\sigma - \sigma_z)}{(j - 1)! \sigma_z} + \frac{(2s)^{j-2}(x - x_z)^{j-2}(\sigma - \sigma_z)^2}{2(j - 2)! \sigma_z^2} \partial_x \sigma + \frac{\alpha^2 (\sigma - \sigma_z)^2}{\nu_z} \partial_x^2 \\
    &= \frac{\rho \alpha}{\nu_z} (\sigma z)^j(x - x_z)^j \frac{1}{j!} + \frac{2s^{j-1}(x - x_z)^{j-1}(\sigma - \sigma_z)}{(j - 1)! \sigma_z} + \frac{s^j(x - x_z)^{j-2}(\sigma - \sigma_z)^2}{(j - 2)! \sigma_z^2} \partial_x \sigma, \quad \text{for } j \geq 3
\end{align*}
\]
where $s = \beta - 1$

Now we assume the implied variance has the expansion:

$$V_{imp} = \tilde{V}(1 + q_1 + q_2 + \cdots),$$

and by comparing the two expressions of $P_n(t)$ and $\tilde{P}_n(t)$, which is given by (??), we get:

$$\tilde{V} = V_z$$

$$q_1 = s(2(x_0 - x_z) - k) + \frac{2(\sigma_0 - \sigma_z)}{\sigma_z} + \frac{\rho \alpha (T - 2k)}{\sqrt{V_z}}$$

$$q_2 = s^2 \left( \frac{T - k^2}{12} - \frac{T^2}{48} \right) + \frac{(\sigma_0 - \sigma_z)^2}{\sigma_z^2} + \frac{\alpha \rho \sigma (\sigma - \sigma_z)(1.5T - k)}{\sigma_z \sqrt{V_z}}$$

$$+ \frac{\alpha^2 (-T^2 + 2T + 4k^2)}{12V_z} + \frac{\alpha^2 \rho^2 (-4T + 5T^2 - 4k^2 - 8Tk)}{16V_z}$$

$$- \frac{s \alpha \rho T (T - 12 - 2k)}{24 \sqrt{V_z}}$$

Recall the CEV expansive coefficients formula (9.10), we notice that when $\alpha = 0$ and $\sigma_z = \sigma_0$, one recovers the CEV coefficients in the first section of this chapter.

Numerical experiment is displayed in Figure 9.2a:

We write our expansive approximation to the implied volatility, in the same notation in the paper [24] again, to show my respect to their ideas:

$$\sigma_B^2 = \frac{\sigma_0^2}{f_{av}^{2-2\beta}} (1 + q_1 + q_2 + \cdots) \tag{9.22}$$

where $f_{av} = \sqrt{F_0 K}$ is the geometric mean of the forward and strike,

$$q_1 = \frac{\alpha \rho (T - 2k) f_{av}^{1-\beta}}{2\sigma_0}$$

$$q_2 = s^2 \left( \frac{T - k^2}{12} - \frac{T^2}{48} \right) + \frac{\alpha^2 (-T^2 + 2T + 4k^2) f_{av}^{2-2\beta}}{12\sigma_0^2}$$

$$+ \frac{\alpha^2 \rho^2 (-4T + 5T^2 - 4k^2 - 8Tk) f_{av}^{2-2\beta}}{16\sigma_0^2} - \frac{s \alpha \rho T (T - 12 - 2k) f_{av}^{1-\beta}}{24 \sigma_0}$$

where $T = \frac{\sigma_0^2}{f_{av}^{2-\beta}}$ is the rescaled time to maturity, $k = \log(F_0/K)$ is the log moneyness, and
Figure 9.2: SABR model European call price, with \( \beta = .85, \alpha = .22, \rho = -.04, \sigma_0 = .35, t = .75 \) for first figure, and \( t = 7.25 \) for the second, \( S_0 = 100 \), with various strikes

\[ s = \beta - 1. \]

or we can write it in the form:

\[ \sigma_B = \frac{\sigma_0}{f_{av}^{1-\beta}}(1 + .5(q_1 + q_2) - .125q_1^2 + \cdots) \]  \hspace{1cm} (9.23)

One can also approximate the marginal pdf for the forward, the methodology is almost the same as the CEV model.

**Theorem 9.2.1.** The marginal pdf for the SABR model will be approximated by:

\[ G_{SABR}(t; F_0, \sigma_0, F_t, \cdot) = \exp(\sigma_z^2 t A) \frac{\exp(- (\log F_0 - \log F_t - \alpha \sigma_z^2 t)^2)}{\sqrt{2\pi t \sigma_z F_t}} \]  \hspace{1cm} (9.24)

where \( \sigma_z = \frac{\sigma_0}{f_{av}^{1-\beta}} \) and \( f_{av} = \sqrt{F_0K} \). \( A \) has the following expansive expression:

\[ A = a_1(\beta - 1) + a_2(\beta - 1)^2 + a_3(\beta - 1)^3 + a_4(\beta - 1)^4 + \cdots \]  \hspace{1cm} (9.25)
where \( s = \beta - 1, T = \frac{\beta t}{\beta^2} \) is the rescaled time to maturity, \( k = \log F_0 - \log K \) is the log-moneyness.

We notice that when \( \alpha = 0 \), we recover the heat kernel approximation formula for CEV in the first section of this chapter.

### 9.3 Approximation to the implied volatility of Heston model-

#### A different approach

Heston model is popular in producing volatility smiles, and for the reason that it is analytically tractable, by Fourier Transform methods. For Heston model, we have the stochastic process for the variance:

\[
dV(t) = \kappa(\theta - V(t))dt + \sigma_V \sqrt{V(t)}dZ(t)
\]

A large amount of literatures are devoted to expand with respect to parameters in the above formula, large \( \kappa \), small \( \sigma_V \), small time to maturity \( t \), small correlation \( \rho \), we also mention that there are literatures deal with the case when \( \rho = 0 \), where some averaging technics were used to get the closed form approximation. However, we take the approach not to specify which parameter is 'small' or 'large', but just keep in mind that \( \kappa \) could be large, \( \sigma_V, \rho \) and time to maturity could be small (but do not have to be), and our expansion will be a generalized version of most of the approximation approaches mentioned above. On the other hand, to prove the convergence of the approximation, we need to specify 'small' or 'large' parameters, such that one can prove the convergence of the expansive series in a rigorous way.
Since $\kappa$ could be "large", we change the variable to get:

$$s(t) = \exp(\kappa t)(V(t) - \theta) - (w_0 - \theta) \tag{9.27}$$

where $w_0$ will be specified later, then we have:

$$\frac{ds(t)}{dt} = \exp(\kappa t)\left[\theta + \exp(-\kappa t)(s(t) + w_0 - \theta)\right] \tag{9.28}$$

Notice that $s(0) = V_0 - w_0$, and that

$$V(t) = \theta + \exp(-\kappa t)(s(t) + w_0 - \theta)$$

Now we write the Feynman-Kac formula for the Heston model in the log price variable $x$ and the above state variable $s$:

$$\partial_t u = \frac{1}{2}(\theta + \exp(-\kappa t)(s + w_0 - \theta))\left[\partial_x^2 u + \frac{1}{2}\sigma^2 V \exp(2\kappa t)(s + w_0 - \theta)\partial_x^2 s\right]$$

$$+ \rho \exp(\kappa t)\sigma V (\theta + \exp(-\kappa t)(s + w_0 - \theta))\partial_x u - ru$$

Notice that we are dealing with parabolic equations with coefficient depend both on time and space. So recall the time-dependent result, Theorem 6.1.1 in chapter 6, and the operators $L_j^t$ are given by the following:

$$L_0^t = \frac{1}{2}(\theta + \exp(-\kappa t)(s + w_0 - \theta))\left[\partial_x^2 u + \frac{1}{2}\sigma^2 V \exp(2\kappa t)(s + w_0 - \theta)\partial_x^2 s\right]$$

$$+ \rho \exp(\kappa t)\sigma V (\theta + \exp(-\kappa t)(s + w_0 - \theta))\partial_x u - ru$$

$$L_1^t = \frac{1}{2}\exp(-\kappa t)(s - s_z)(\partial_x^2 - \partial_x)u + \frac{1}{2}\sigma^2 V \exp(\kappa t)\partial_x^2 s + \rho \sigma V (s - s_z)\partial_x s$$

$$L_j^t = 0, \text{ for } j \geq 2$$

The special structure of $L_j^t$ makes the computation of $\mathcal{P}(t)$ extremely simple, especially for the first few terms, we get by Theorem 6.1.1:

$$\mathcal{P}_1(t) = \frac{1 - \exp(-\kappa t)}{2\kappa}(s - s_z)(\partial_x^2 - \partial_x)u + \frac{\rho \sigma V}{2\kappa^2}(f(\kappa t)(s + w_0 - \theta) + g(\kappa t)\theta)(\partial_x^3 - \partial_x^2 s)$$

$$\mathcal{P}_2(t) = ...$$

where $f(t) = \exp(-t) - 1 + t$, and $g(t) = \exp(t) - 1 - t$
Recall that the solution is:

$$u(t, x, s) = \sum_{n=0}^{\infty} P_n(t) u^0(t, x, s)$$  \hspace{1cm} (9.29)

and that $u^0$ solves the equation:

$$\partial_t u^0(t; x, s) = L_0^t u^0(t; x, s)$$  \hspace{1cm} (9.30)

We find that

$$u^0(t; x, s) = \text{Call}_{BS}(S = \exp(x), \sigma = \sqrt{V_z}, K, r, t)$$  \hspace{1cm} (9.31)

where $V_z$ is the averaged volatility of the Black-Scholes formula with deterministic time-dependent volatilities:

$$V_z = \frac{1}{t} \int_0^t (\theta + \exp(-\kappa \tau)(s_z + w_0 - \theta))d\tau$$

$$= \theta + \frac{1 - \exp(-\kappa t)}{\kappa t}(s_z + w_0 - \theta)$$

Now we following the procedure of finding the implied volatility, which was described in the first section of this chapter, to assume:

$$V_{imp} = q_0 + q_1 + \cdots$$  \hspace{1cm} (9.32)

Then by comparing $P_0(t)$, we get:

$$q_0 = V_z,$$

Nontrivial computation shows that by comparing $P_1(t)$, one gets:

$$q_1 = 1 - \exp(-\kappa t)\left(V_0 - m_0\right) + \rho \sigma V \left(\frac{V_z t - 2k}{2 \kappa^2 t^2 V_z}\right)\left((f(\kappa t)(m_0 - \theta) + g(\kappa t)\theta\right)$$  \hspace{1cm} (9.33)

where $m_0 = s_z + w_0$ and $k = \log(S_0) - \log(K) + rt$ is the log-moneyness

Note that one is free to choose $m_0$, a naive choice is $m_0 = V_0$, there are other possible choices. Numerical experiment is given in Figure 9.3b, and $m_0$ is taken to $V_0$. 

96
Figure 9.3: Heston model European call price, with $t = .25$, $\kappa = 5$, $\theta = .2$, $\sigma_V = .26$, $\rho = - .01$, $V_0 = .14$, $S_0 = 50$, $r = .05$, and various strikes.
Chapter 10  
Conclusion and Future Research Directions

In this dissertation, we further developed the Dyson-Taylor commutator method. For the theoretical part, we obtained the explicitly the express for $P_n(t)$ in Theorem 2.2.3, which is important to get the properties of the approximating series. We also obtained the inductive relationship of $P_n(t)$ in (2.2.1), which is essential in computation. This enables us to explicitly compute the higher order approximation to the option price/transition probability density/implied volatility, both numerically and symbolically. Specifically, we examined CEV model, SABR model and Heston model in detail, but also established the general framework of approximating general diffusion based models. Time dependent coefficients PDE and partial integral differential equations resulting from jump-diffusion model are also explored.

We think there are several direction which deserve further exploration:

1. There have been several papers, which applied the Malliavin calculus to approximate the jump-diffusion models [43–47]. We believe that essentially the results should also be obtained by PDE approach. However, I do have difficulties when computing the commutator of a second order partial differential operator and an integral operator. It seems that those who use SDE approaches do not have such problems. Taking into consideration the popularity of jump-diffusion models, and the fact that most of the jump-diffusion models rely on Fourier Transform methods to get a solution, it is worth for us to understand the Malliavin calculus approach, and eventually obtain more results for jump models.

2. Our methods are from the perspective of PDE. Because of the Feynman-Kac formula,
it could be used in SDE as well. For example, in the numerical simulation of the SDE:

\[ dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t \quad (10.1) \]

with \( X_0 = x_0 \), where \( W_t \) is a standard Brownian motion. For some \( 1 < M < 2 \) A plausible Monte Carlo simulation of the sample path, based on an acceptance-rejection method, could be:

1. For small time interval \( \Delta t \), one generate \( y = x_0 + \mu(z_0, 0)\Delta t + \sigma(z_0, 0)N(0, \Delta t) \)
2. Compute the transition probability density \( P(t; x_0, y) = A_nP^0(t; x_0, y) \), by the approximated expansion, truncated to some high order term, where \( P^0(t; x_0, y) \) is the transition probability density for the process:

\[ dX_t = \mu(z_0, t)dt + \sigma(z_0, t)dW_t \quad (10.2) \]

3. if \( A_n < M \), reject the updating \( x_0 \rightarrow y \); otherwise, accept the updating. Go to step 1 until we reach the time level we are concerned.

It is expected that when \( \Delta t \) is sufficiently small, \( M \) could be very closed to 1, which will make the simulation effective. On the other hand, it can be expected that the simulation has high order convergent rate, given the heat kernel is approximated to high orders. Last, in this approach, we avoided to get complicated algorithms, which always involve Brownian bridge, in traditional numerical stochastic differential equation. See chapter 10 in [65], for example.

3. The geometric approach is very attractive, in that if one could find the numerical method to compute higher order terms. In [21], the author cited a expansion formula for heat kernel on manifold (Theorem 3.1), and mentioned that the expansive coefficients \( a_i(x, y) \) has recently(2005) been computed up to order \( n = 8 \). However, the author does not give explicitly the references further. It could be expected that the numerical algorithm developed in this paper has the chance to be used to compute those coefficient, which are purely ‘geometric’ variates, meaning that they are independent of the choice of the coordinate systems, such as the scalar curvature \( R \), the Christoffel’s coefficients \( \Gamma^i_{jk} \), the Riemann tensor \( R_{ijkl} \), and the Ricci tensor \( R_{ij} \).

4. Although we focus on the equity and foreign exchange derivative models in this paper, there are no reason to stop us to apply this methodology to credit derivatives like Credit Default Swaps(CDS), and Credit Default Swaptions, where the implied default rate played
the same role of implied volatility in equity/foreign exchange options. I will continue in this direction after I work in the credit related group.

5. I will further develop the MATLAB code, and possible rewrite all the codes in C++ to improve the efficiency. The final product is expected to be something, when the user input the operators $L_j$ in (4.11), the program will automatically produce the coefficients in the approximated heat kernel in equations and the coefficients in the approximated implied volatility in equations.

6. We plan to plant the approximated pricers we obtained in this paper (CEV, SABR, Heston, local volatility Kou, local volatility Bates models...) into a statistical inference framework, such that we can fast calibrate the historical/real time market data. This is an ongoing project with Professor Liechty and Professor Farnsworth.

7. Although the numerical results seems good, it is still necessary to prove the convenience of the expansion approximation series. It can be noticed that in most of the literatures, there is a specified small parameter (small correlation, small volvol, small (1-correlation) ...), but we did not explicitly state which parameter is small, but just treat the series expansion as a formal asymptotic expansion. The advantage is that I do not to concern with respect to which parameter we expand the expression, which enable us to tackle more complicated models. On the other hand, it adds difficulty to the proof of convergence. We will treat this problem later.
Appendix A
Computing Operators Multiplications and Commutators

In this section, we will introduce the ideas in computational aspects throughout this dissertation. All computational are coded in MATLAB. In the first section, it is shown how differential operators in $D(a, b)$ are stored, for both one dimensional and higher dimensions. The we will illustrate how operator multiplications and commutators are computed, by a precomputed 6-dimensional matrix $S$, which is the multiplication of elementary operators, see Definition A.1.1 below. In the second section, I will illustrate the MATLAB code to compute $P_n(t)$, by the inductive relationship in Lemma 2.2.1. This is the critical part for all this paper, since one of the contributions of this paper is that I can computed arbitrary higher order approximations by programming, either by numerical or symbolical computation. In the third section, The details of how we got the approximated option price/heat kernel/implied volatility were obtained, after we obtained $P_n(t)$.

A.1 Representation of differential operator by matrix and compute their multiplication

A.1.1 One dimensional operator multiplication and commutator computation

In this section, $L \in D(a, b)$, which is defined in chapter 2, and $A$ is always a matrix.
Definition A.1.1. An elementary operator is an differential operator in the form of

\[ E_{i,j} = x^{j-1} \frac{\partial^{i-1}}{\partial x^{i-1}}. \]

An elementary matrix is a matrix with all entries 0, except one entry which equals to 1. We see that there is a one-to-one correspondence between elementary matrix and elementary operator. Furthermore, there is a one-to-one correspondence between a matrix \( A \) and an operator \( L_A \):

\[ L_A = \sum_{i,j=1} A(i, j) x^{j-1} \frac{\partial^{i-1}}{\partial x^{i-1}} \] (A.1)

and if \( L = \sum_{i,j=1} L_{i,j} x^{j-1} \frac{\partial^{i-1}}{\partial x^{i-1}} \)

\[ A_L(i, j) = L_{i,j} \] (A.2)

A 6-dimensional matrix is precomputed to store the multiplication of elementary operators:

\[ S(:, :, i, j, k, l) = A_{E_{i,j}E_{k,l}} \] (A.3)

The following formula:

\[ \frac{\partial^n}{\partial x^m} \left( x^n \frac{\partial^p}{\partial x^p} \right) = \frac{\partial^{n-1}}{\partial x^{m-1}} \left( n x^{n-1} \frac{\partial^p}{\partial x^p} + x^n \frac{\partial^{p+1}}{\partial x^{p+1}} \right) \] (A.4)

enables us to compute \( S \) by induction, with respect to the order of differential \( m \). Figure A.1 shows part of the MATLAB code to compute \( S \).

Next, operator multiplication will be handled by decomposition into elementary operators and using the matrix \( S \): For any two operators \( L^1 \) and \( L^2 \):

\[ L^1 = \sum_{i,j} L^1_{i,j} x^{j-1} \frac{\partial^{i-1}}{\partial x^{i-1}} \]

\[ L^2 = \sum_{k,l} L^2_{k,l} x^{l-1} \frac{\partial^{k-1}}{\partial x^{k-1}} \]
Figure A.1: Code to compute $S$

The multiplication of $L^1$ and $L^2$ is given by:

$$A_{L^1L^2} = \sum_{i,j,k,l} L^1_{i,j} L^2_{k,l} S(; ; i, j, k, l)$$  \hspace{1cm} (A.5)

The commutator is computed by:

$$[L^1, L^2] = L^1 L^2 - L^2 L^1$$  \hspace{1cm} (A.6)

A.1.2 Two and higher dimensions

The idea is that I still use a matrix to store $L$:

$$x^{j_1-1} y^{j_2-1} \frac{\partial^{i_1-1}}{\partial x^{i_1-1}} \frac{\partial^{i_2-1}}{\partial y^{j_2-1}}$$

In order to do so, we used the one-to-one map $f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$, given by:

$$i = f(i_1, i_2) = \frac{(i_1 + i_2)(i_1 + i_2 - 1)}{2} - i_2 + 1$$  \hspace{1cm} (A.7)

Given

$$L = \sum_{i_1,j_1,j_2,j_2} L_{i_1,j_1,i_2,j_2} x^{j_1-1} y^{j_2-1} \frac{\partial^{i_1-1}}{\partial x^{i_1-1}} \frac{\partial^{i_2-1}}{\partial y^{j_2-1}}$$
Then $A_L$ is matrix to represent this operator with:

$$A_L(f(i_1, i_2), f(j_1, j_2)) = L_{i_1, j_1, i_2, j_2}$$  \(\text{(A.8)}\)

We do not explicitly write out the other direct:

$$A \rightarrow L_A,$$

but keep in mind it is well defined.

For two differential operators:

$$L^1 = \sum_{i_1, j_1, i_2, j_2} L^1_{i_1, j_1, i_2, j_2} x^{j_1-1} y^{j_2-1} \frac{\partial^{i_1-1}}{\partial x^{i_1-1}} \frac{\partial^{i_2-1}}{\partial y^{i_2-1}}$$

$$L^2 = \sum_{k_1, l_1, k_2, l_2} L^2_{k_1, l_1, k_2, l_2} x^{l_1-1} y^{l_2-1} \frac{\partial^{k_1-1}}{\partial x^{k_1-1}} \frac{\partial^{k_2-1}}{\partial y^{k_2-1}}$$

The multiplication of $L^1$ and $L^2$ is given by:

$$A_{L^1 L^2} = \sum_{i_1, j_1, i_2, j_2, k_1, l_1, k_2, l_2} L^1_{i_1, j_1, i_2, j_2} L^2_{k_1, l_1, k_2, l_2} S(:, :, i_1, j_1, i_2, j_2) \otimes S(:, :, k_1, l_1, k_2, l_2)$$  \(\text{(A.9)}\)

where the linear operator $\otimes$, of two elementary matrix $A, B$ with $A(i, j) = 1$ and $B(k, l) = 1$, is defined as an elementary matrix

$$C = A \otimes B$$  \(\text{(A.10)}\)

with $C(f(i, k), f(j, l)) = 1$.

The commutator is computed by:

$$[L^1, L^2] = L^1 L^2 - L^2 L^1$$  \(\text{(A.11)}\)

One realizes that it is straightforward to generalize the approach above to encompass higher order operators, by define a one-to-one map $f : \mathbb{N}^3 \rightarrow \mathbb{N}$. 

104
A.2 Compute $\mathcal{P}_n(t)$

A.2.1 Compute $\mathcal{P}_n(t)$ inductively

As far as computation of the operator multiplication/commutator of differential operators in $\mathcal{D}(a,b)$ are handled, $\mathcal{P}_n(t)$ can be computed. We have inductive representation of $\mathcal{P}_n(t)$ by Lemma 2.2.1, or the explicit expression by Theorem 2.2.3. The explicit expression has its advantage, especially for theoretical analysis. However, from the algorithm point of view, the former is more effective, because of its natural bottom-up memorization structure. Please see [74], Chapter 15 for a detailed discussion. So we proceed to compute inductively $\mathcal{P}_n(t)$ by Lemma 2.2.1. We realized that Lemma 3.1.8 is not only of theoretical importance, but is also essential for the numerical/symbolic computation. Specifically, I prove the following lemmas, which is more closely related to the practical issue:

Lemma A.2.1. The range of order of time variable $t$ in $\mathcal{P}_n(t)$ ranges from 2 to $2n$.

Proof. This is from the explicit expression of $\mathcal{P}_n(t)$ in Theorem 2.2.3

From Lemma 3.1.8 and Lemma A.2.1, memories are only assign for differential operators up to order $3n$, polynomials up to order $n$, and order of $t$ to $n$, in computing $\mathcal{P}_n(t)$. Such that we discard or other terms, and know where to stop, such that we can efficiently compute $\mathcal{P}_n(t)$, Figure A.2 shows the key part of MATLAB code computing $\mathcal{P}_n(t)$, by the inductive relationship in Lemma 2.2.1

A.2.2 Efficient way of computing for two and higher dimensions

In practice, the computation of $\mathcal{P}_n(t)$ for two and higher order terms are time consuming, especially when we carry out symbolic computations. Let us write $x \in \mathbb{R}^N$ as $x = (x_1, \tilde{x})$. One notices the special structure of the higher dimension solution $u^0(t,x)$, that it only depends on the first coordinate $x_1$. The feature makes it possible to make the computation of $\mathcal{P}_n(t)$ faster, by the following lemma:

Lemma A.2.2. When all terms containing differentials to $\tilde{x}$ are discarded in computing $\mathcal{P}_n(t)$, by the inductive relation Lemma 2.2.1, one gets the same approximating result (4.12) as if all terms are kept.

Proof. The first observation is that $u^0(t,x)$ only depends on time and the first space variable, such that any differential operator containing differentials to $\tilde{x}$, when acting on $u^0(t,x)$, gives
Figure A.2: Matlab code to compute $P_n(t)$. The number of commutator computations are accurately controlled, by the order estimates Lemma 3.1.8 and time variable estimation Lemma A.2.1

zero. The second observation is that when there are differentials to $\tilde{x}$, these differentials remain under operator multiplication from left, and commutators with constant-coefficient differential operators, which appears in the inductive relationship 2.2.1. The lemma is proved when one combines these two observations.

Figure A.3 shows how we drop those terms with differentials of $\tilde{x}$, for two dimensional case.

### A.3 Compute closed form approximation to option price/heat kernel/implied volatility

Recall the expansion formula (4.12). After I obtained $P_n(t)$ to certain order $m$, numerical or symbolical computation gives back the corresponding closed form approximation to option price, heat kernel and implied volatilities. As I discussed in the introduction, these three terms are closely related, but they have not to be consequence from each other. From the computational aspect, I treat them differently, in some sense separately. I first show how
Figure A.3: Matlab code to compute two dimensional operator multiplication. Several issues are addressed: accurate order estimate such that I discard higher order terms, discard terms containing differentials of $\tilde{x}$, decompose the operator into elementary operator multiplication, which used saved value, and use the one-to-one correspondence $f : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ to save two dimensional differential operates in matrix

to get the implied volatility expansion in detail, and the other two are obtained in slightly different ways: I get the closed form approximation to option price directly from the expansibe formulae (4.12). The heat kernel expansion is obtained in the form of:

$$G(t; x, y) = AG^0(t; x, y)$$  \hspace{1cm} (A.12)

where the term $A$ will be explained in detail later.

### A.3.1 computation of the implied volatility

To approximate the implied volatility, I consider the following Black-Scholes-like formulae:

$$\partial_t u = \frac{V_{imp}}{2} (\partial_x^2 - \partial_x)u + r\partial_x u - ru$$  \hspace{1cm} (A.13)

where $V_{imp}$ is the implied volatility, with expansive approximation:

$$V_{imp} = \sum_{j=0}^{\infty} V_j$$  \hspace{1cm} (A.14)
The next step is crucial, which is different from all literature within my knowledge. Instead of obtaining the Black-Scholes formula in closed form, I solve (A.13) by the expansive formula (4.12), to get:

\[ u(t, x) = \sum_{j=0}^{\infty} \tilde{P}_j(t) u^0(t, x) \]  
(A.15)

where \( \tilde{P}_j(t) \) is given inductively by Lemma 2.2.1, and that \( \tilde{L}^j \) are given by:

\[ \tilde{L}^0 = \frac{V_0}{2}(\partial^2_x - \partial_x) + r\partial_x - r \]  
(A.16)

\[ \tilde{L}^j = \frac{V_j}{2}(\partial^2_x - \partial_x) \quad \text{for } j \geq 1 \]  
(A.17)

Now I proceed to compare the two expansive expressions to get:

\[ \tilde{P}_j(t)\tilde{u}^0(t, x) = P_j(t)u^0(t, x) \]  
(A.18)

where \( P_j(t) \) is computed from the models we are concerning. Also notice that \( \tilde{u}^0(t, x) \) solves

\[ \partial_t \tilde{u}^0 = \frac{V_0}{2}(\partial^2_x - \partial_x)\tilde{u}^0 + r\partial_x\tilde{u}^0 - r\tilde{u}^0 \]  
(A.19)

Taking European Call option for example.

\[ \tilde{u}^0(t, x) = e^{x \tilde{d}_1} N(\tilde{d}_1) - Ke^{-rt} N(\tilde{d}_2) \]

where \( \tilde{d}_1 = \frac{x - \log K + (r + .5V_0)t}{\sqrt{V_0t}} \) and \( \tilde{d}_2 = d1 - \sqrt{V_0t} \).

\( u^0(t, x) \) also satisfies a Black-Scholes formula:

\[ \partial_t u^0 = L_0 u^0 \]  
(A.20)

when we choose certain operator \( L_0 \), such that the solution is also in Black-Scholes form:

\[ u^0(t, x) = e^{x d_1} N(d_1) - Ke^{-rt} N(d_2) \]

where \( d_1 = \frac{x - \log K + (r + .5V_z)t}{\sqrt{V_zt}} \) and \( d_2 = d1 - \sqrt{V_zt} \). \( V_z \) is related to the specific model and also related to the point \( z \) where space variables are frozen. For example, in the CEV model 1.4, \( V_z = \sigma^2 \exp(2(\beta - 1)z) \).
From Lemma 2.2.1 and (A.18):

\[
\left( \tilde{L}n + \sum_{l=1}^{n-1} \int_0^t e^{(t-\tau)ad_{L_0}} (\tilde{L}_l \tilde{P}_{n-l}(\tau)) d\tau \right) \tilde{u}^0(t, x) = \mathcal{P}_n(t) \tilde{u}^0(t, x)
\] (A.21)

We get for \( n \geq 0 \)

\[
\tilde{L}_n \tilde{u}^0(t, x) = \mathcal{P}_n(t) \tilde{u}^0(t, x) - \left( \sum_{l=1}^{n-1} \int_0^t \tilde{L}_l \tilde{P}_{n-l}(\tau) d\tau \right) \tilde{u}^0(t, x)
\]

\[
\tilde{P}_n(t) = \sum_{l=1}^{n} \int_0^t \tilde{L}_l \tilde{P}_{n-l}(\tau) d\tau
\]

Which can be solved inductively. Notice that since \( \tilde{L}_j \) are constant coefficients, there is no commutator computation needed.

We observe from the comparing the leading terms, one can choose \( V_0 = V_z \), and thus \( \tilde{u}^0(t, x) = u^0(t, x) \)

Recall (A.16), we have for \( n \geq 1 \):

\[
\tilde{L}_j \tilde{u}^0(t, x) = \frac{V_j e^{x p(d_1)}}{2 \sqrt{V_0 t}}
\] (A.22)

where \( p(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \).

I have the following algorithm, to compute \( V_j \) up to order \( m \):

1. Compute \( \mathcal{P}_n(t) \) for \( 1 \leq n \leq m \), inductively by Lemma 2.2.1
2. \( V_0 = V_z, \tilde{P}_0(t) = Id \)
3. let \( j = 1 \)
4. compute
   \[
   V_j = \frac{2 \sqrt{V_0} \left( \mathcal{P}_j(t) - \sum_{l=1}^{j-1} \int_0^t \tilde{L}_l \tilde{P}_{j-l}(\tau) d\tau \right) \tilde{u}^0(t, x)}{e^x p(d_1) \sqrt{t}}
   \] (A.23)
5. Define \( \tilde{L}_j = \frac{V_j}{2} (\partial_x^2 - \partial_x) \), and compute

\[
\tilde{P}_j(t) = \sum_{l=1}^{j} \int_0^t \tilde{L}_l \tilde{P}_{j-l}(\tau) d\tau = \sum_{l=1}^{j-1} \int_0^t \tilde{L}_l \tilde{P}_{j-l}(\tau) d\tau + \tilde{L}_j t
\] (A.24)
6 if \( j \leq m \), go to 3; else, stop and return \( V_n \), for \( 0 \leq n \leq m \)

A.3.1.1 Compute higher order derivatives - Hermite polynomials

One still needs to compute terms in the form of

\[
P_n(t)u^0(t, x) \tag{A.25}
\]

where \( P_n(t) \in D(n, 3n) \), \( u^0(t, x) \) is the Black-Scholes solution given by:

\[
u^0(t, x) = e^x N(d_1) - Ke^{-rt} N(d_2)
\]

where \( d_1 = \frac{x - \log K + (r + .5V_0)t}{\sqrt{V_0}t} \) and \( d_2 = d_1 - \sqrt{V_0}t \), and \( N(x) \) is the cumulative distribution function of standard normal distribution, which is given by:

\[
N(x) = \int_{-\infty}^x p(y)dy \tag{A.26}
\]

where \( p(y) = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \). I explicitly compute the higher order derivatives of \( \frac{\partial^n u^0(t, x)}{\partial x^n} \) for \( n \geq 1 \): \( \frac{\partial u^0(t, x)}{\partial x} = e^x N(d_1) \), which from the computation of the hedging ratio \( \Delta \) [75].

For \( n \geq 1 \)

\[
\frac{\partial^{n+1} u^0(t, x)}{\partial x^{n+1}} = \frac{\partial^n}{\partial x^n} (e^x N(d_1))
\]

\[
= e^x \sum_{j=0}^{n} \binom{n}{j} \frac{\partial^j}{\partial x^j} N(d_1)
\]

\[
= e^x \left( N(d_1) + p(d_1) \sum_{j=0}^{n-1} (-1)^j \binom{n}{j+1} \frac{H_j(d_1)}{(V_0 t)^{\frac{j+1}{2}}} \right)
\]

where \( H_j(x) \) is the Hermite polynomial, which is given by:

\[
H_0(x) = 1
\]
\[ H_1(x) = x \]
\[ H_n(x) = xH_{n-1}(x) - (n-1)H_{n-2}(x), \quad \text{for } n \geq 2 \]

We used the following facts in the computations above:

\[ \frac{\partial N(x)}{\partial x} = p(x), \]
\[ \frac{\partial^j}{\partial x^j} p(x) = (-1)^j H_j(x)p(x), \]

and

\[ \frac{\partial d_1}{\partial x} = \frac{1}{\sqrt{V_0 t}} \]

### A.3.2 Algorithm to compute the closed form approximation to option price

For a model whose \( P_n(t) \) has been computed, we can expression the European call option price in the form of:

\[ C = BS + \sum_{j=1}^{\infty} c_j \quad (A.27) \]

The algorithm to compute \( H_j \) is given below:

1. Load computed/Compute \( P_n(t) \)
2. for \( j \) from 1 to \( m \)
3. compute

\[ c_j = P_j(t)u^0(t,x) \quad (A.28) \]
A.3.3 Algorithm to compute the closed form approximation to heat kernel

In order to approximate the heat kernel, given the computed $P_n(t)$ for a given model, we write a related Black-Scholes like formula as:

$$\partial_t u = \frac{V_0}{2} (\partial_x^2 - \partial_x) u + r \partial_x u - ru + \left( \sum_{j=1}^{\infty} A_j \right) u$$  \hspace{1cm} (A.29)

Such that the heat kernel is given by:

$$G(t, x, y) = G^0(t; x, y) \exp(t \sum_{j=1}^{\infty} A_j)$$  \hspace{1cm} (A.30)

where

$$G^0(t; x, y) = \frac{\exp(- \frac{(x - y + (r - .5V_0)t)^2}{2V_0t})}{\sqrt{2\pi V_0t}}$$

We follow the same idea of computing the volatility. The difference here is that we assume another Black-Scholes like expression, such that we got more clean results. The following algorithm produces $A_j$ inductively:

I have the following algorithm, to compute $V_j$ up to order $m$:

1. Load computed/Compute $P_n(t)$
2. $V_0 = V_{\tilde{z}}$, $\tilde{P}_0(t) = Id$
3. let $j = 1$
4. compute

$$A_j = \left( \frac{P_j(t) - \sum_{l=1}^{j-1} \int_0^t \tilde{L}_l \tilde{P}_{j-l}(\tau)d\tau}{G^0(t; x, y)} \right) \frac{G^0(t; x, y)}{t}$$  \hspace{1cm} (A.31)

5. Define $\tilde{L}_j = A(j)$, which is constant multiplication operator, and compute

$$\tilde{P}_j(t) = \sum_{l=1}^{j} \int_0^t \tilde{L}_l \tilde{P}_{j-l}(\tau)d\tau = \sum_{l=1}^{j-1} \int_0^t \tilde{L}_l \tilde{P}_{j-l}(\tau)d\tau + \tilde{L}_j t$$  \hspace{1cm} (A.32)

6. if $j \leq m$, $j = j + 1$, go to 3; else, stop and return $A_n$, for $0 \leq n \leq m$

The codes for computing option price, heat kernel and implied volatility were written...
altogether, and is show in Figures A.4, A.5 and A.6.
Figure A.6: CEV model for $\alpha = .85, S_0 = 50, \sigma = .34, t = .25, r = .00$
Bibliography


[33] Kampen, J., *How to compute the length of a geodesic on a Riemannian manifold with small error in arbitrary Sobolev norms,* arXiv:0811.1734v1,


Vita
Chao Liang

Chao Liang was born in 1984 in Sichuan Province, the People’s Republic of China. In 2000, he enrolled in the Shuguang High School in the city of Zigong. There he was trained in science major. In 2003, he was admitted by the Department of Mathematics, Tsinghua University. He obtained the Bachelor’s and Master’s degree in Pure and Applied Mathematics from Tsinghua University, in 2007 and 2009 respectively. Then he was accepted by the PhD program of applied mathematics at Penn State University, where his research was focused on numerical/analytical approximations to PDE, with application to financial modeling. He will graduate on August 2014, and become an Associate at Fixed Income Division, Morgan Stanley.