SEMIMARTINGALES, MARKOV PROCESSES AND THEIR APPLICATIONS IN MATHEMATICAL FINANCE

BY JIN WANG

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Paul Feehan

and approved by

New Brunswick, New Jersey

October, 2010

ABSTRACT OF THE DISSERTATION

Semimartingales, Markov processes and their Applications in Mathematical Finance

by Jin Wang

Dissertation Director: Paul Feehan

We show that the marginal distribution of a semimartingale can be matched by a Markov process. This is an extension of Gyöngy's theorem to discontinuous semimartingale.

Acknowledgements

I would like to thank my advisor, Professor Paul Feehan. His support and kindness helped me tremendously. I am deeply grateful for his constant encouragement and insightful discussions in all these years at Rutgers.

I would like to thank all of my committee members, Professor Dan Ocone, Professor Richard Gundy and Professor Panagiota Daskalopoulos, for their comments and encouragements.

I also want to thank the mathematics department of Rutgers for its support. Last but not least, I want to thank all my friends at Rutgers, they made my PhD life colorful and enjoyable.

Dedication

To my parents, Wang Xizhi, Wu Jinfeng.

Table of Contents

Chapter 1

Introduction

In order to price and hedge financial derivatives, stochastic process models of the dynamics of the underlying stocks have been introduced. The Black-Scholes model is based on the assumption that the stock price process follows a geometric Brownian motion with constant drift and volatility. It is well known that this model is too simple to capture the risk-neutral dynamics of many price processes. Dupire [10] significantly improved the Black-Scholes model by replacing the constant volatility parameter by a deterministic function of time and the stock price process. By construction, the asset price process given by Dupire's local volatility model has the same one-dimensional marginal distributions as the market price process. Therefore, European-style vanilla options whose values are determined by the marginal distributions can be priced correctly. In practice, the local volatility model is widely used to price not only vanilla options, but also complex options with path dependent payoffs, even though such options cannot be priced correctly by Dupire's model.

Dupire's model is a practical implementation of a famous result of Gyöngy [15]. Given an Itô process, Gyöngy proved the existence of a Markov process with the same one-dimensional marginal distributions as the given Itô process. The coefficients defined in Gyöngy's process are given by Dupire's local volatility model.

In [8] Brunick generalized Gyöngy's result under a weaker assumption so that the prices of path dependent options can be determined exactly. For example, consider a barrier option whose value depends on the joint distribution of the stock price and its running maximum; Brunick's result shows that there exists a two-dimensional Markov process with the same joint distributions. It gives a model process which is perfectly calibrated against market data, but simpler than the market process.

Bentata and Cont [3], [4] extended Gyöngy's theorem to semimartingales with jumps. They showed that the flow of marginal distributions of a discontinuous semimartingale could be matched by the marginal distributions of a Markov process. The Markov process was constructed as a solution to a martingale problem. However, their proof of the main theorem is incomplete. Using this result, they derived a partial integro-differential equation for call options. This is a generalization of Dupire's local volatility formula. We will discuss the details in Section 2.2.3.

Motivated by these results, we present a partial differential equation based proof to the mimicking theorem of semimartingales in this thesis.

The thesis is organized as follows:

In chapter 2, we review some stochastic process models and mimicking theorems. The first section is dedicated to volatility models: the Black-Scholes model, stochastic volatility models, including Heston model and SABR model, Dupire's local volatility model, and some jump models, including Merton's jump diffusion model and Bates' model. In the second section we introduce the theorems of Gyöngy, Brunick, and Bentata and Cont.

In chapter 3, we give a new proofs of Gyöngy's theorem. The proof is based on a uniqueness result of a parabolic equation. We first derive the partial differential equation satisfied by the marginal distributions of an Itô process. Then we construct a Markov process and show that the marginal distributions of this Markov process satisfy the same partial differential equation. By the uniqueness result in [14], we conclude that the Markov process has the same marginal distributions as the Itô process.

In chapter 4, we extend the partial differential equation proof of Gyöngy's theorem for Itô process to the case of semimartingale processes. We first derive the forward equation satisfied by the probability density functions of a semimartingale and show it is obeyed by the probability density functions of the mimicking process. This forward equation is a partial integro-differential equation (PIDE). By analogy with the construction of fundamental solutions to a partial differential equation, we construct fundamental solutions of the PIDE using the parametrix method [14]. Through our construction, we discover the conditions which ensure this equation has fundamental solutions. These conditions guarantee the existence of the transition probability density function for the semimartingales. Then we derive and apply energy estimates to prove uniqueness of the fundamental solution.

Pseudo-differential operators should be natural tools to study our partial integro-differential equation. In chapter 5, we recall the relevant theory of pseudo-differential operators, discuss their relationship with the martingale problem and indicate areas of further research.

In the last chapter, we apply Brunick's result and derive a local volatility formula for single barrier options. The exotic options market is most developed in the foreign exchange market. This formula allows us to price exotic options in the FX market. And these prices are consistent with the market prices of single barrier options we observe in the market.

More ideas for further research

1. In Gyöngy's theorem, the drift and covariance processes are bounded, and the covariance process satisfies the uniformly elliptic condition. However, the covariance process in the Heston model is a CIR process which is neither bounded nor bounded away from zero. So we plan to remove these constraints.

2. We plan to give a numerical illustration of Cont-Bentata's result using Monte Carlo simulation.

3. We also plan to give a numerical illustration of Brunick's result by computation of up-and-out call prices for $S(0) \in [0, S_{max}]$ using original and mimicking processes, where mimicking processes are geometric Brownian motion or Heston process.

Chapter 2

Background

In this chapter, we review some models and theorems. The first section is dedicated to volatility models: the Black-Scholes model, stochastic volatility models, Dupire's local volatility model, and some jump models. In the second section we introduce Gyöngy's theorem, Brunick's theorem, and Bentata and Cont's result.

2.1 Volatility models

2.1.1 The Black-Scholes model

The Black-Scholes model [5] shows how to price options on a stock. An European call option gives the right to its owner to buy at time T one unit of the stock at the price K, where T is called date of maturity and a positive number K is called strike or exercise price. The well known Black-Scholes formula gives the price of such an option as a function of the stock price S_0, the strike price K, the maturity T, the short rate of interest r, and the volatility of the stock σ. Only this last parameter is not directly observable but it can be estimated from historical data. Call options are now actively traded. The work of Black and Scholes on how to price call options seems irrelevant as the prices are already known. However, this is not the case. First, there exist more complex options, often called exotic options, a simple example that we will consider later is a barrier option. These are not actively traded and therefore need to be priced. More importantly the work of Black and Scholes shows that "it is possible to create a hedged position, consisting of a long position in the stock and a short position in [calls on the same stock], whose value will not depend on the price of the stock". In other words, they provide a method to hedge risk of a portfolio in the future. The replicating portfolios have to be rebalanced continuously in a precise way. The Black-Scholes pricing is widely used in practice, because it is easy to calculate and provides an explicit formula of all the variables.

However, prices obtained from the Black-Scholes formula are often differernt from the market prices. In the model, the underlying asset is assumed to follow a geometric Brownian motion with constant volatility σ. That is,

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

where $W(t)$ is a Brownian motion. Given the market price of a call or put option, the implied volatility is the unique volatility parameter to be put into the Black-Scholes formula to give the same price as the market price. At a given maturity, options with different strikes have different implied volatilities. When plotted against strikes, implied volatilities exhibit a smile or a skew effect. Although volatility is not constant, results from the Black-Scholes model are helpful in practice. The language of implied volatility is a useful alternative to market prices. It gives a metric by which option prices can be compared across different strikes, maturities, underlying, and observation times.

2.1.2 Stochastic volatility models

Stochastic volatility models are useful because they explain in a self-consistent way why options with different strikes and expirations have different Black-Scholes implied volatility. Under the assumption that the volatility of the underlying price is a stochastic process rather than a constant, they have more realistic dynamics for the underlying.

Suppose that the underlying price S and its variance V satisfy the following SDEs,

$$\begin{aligned}
dS_t &= \mu_t S_t dt + \sqrt{V_t} S_t dW_1, \\
dV_t &= \alpha(S_t, V_t, t)dt + \eta\beta(S_t, V_t, t)\sqrt{V_t} dW_2, \\
&< dW_1, dW_2 >= \rho dt,
\end{aligned}$$

where μ_t is the instantaneous drift of S, η is the volatility of the volatility process and ρ is the correlation between random stock price returns and changes in V_t, W_1, W_2 are Brownian motions.

The stochastic process followed by the stock price is equivalent to the one assumed in the Black-Scholes model. As η approaches 0, the stochastic volatility model becomes a time-dependent volatility version of the Black-Scholes model. The stochastic process followed by the variance can be very general. In the Black-Scholes model, there is only one source of randomness, the stock price, which can be hedged by stocks. In the stochastic volatility model, random changes in volatility also need to be hedged in order to construct a risk free portfolio.

The most popular and commonly used stochastic volatility models are the Heston model and the SABR model.

In the Heston model [16], the volatility follows a square root process, namely a CIR process

$$dV_t = \theta(\omega - V_t)dt + \eta\sqrt{V_t}dW_2, \tag{2.1}$$

where ω is the mean long-term volatility, θ is the speed at which the volatility reverts to its long-term mean and η is the volatility of the volatility process.

the SABR model (stochastic alpha, beta, rho)[22] describes a single forward F under stochastic volatility V_t:

$$\begin{aligned} dF_t &= V_t F_t^{\beta}\, dW_1, \\ dV_t &= \alpha V_t\, dW_2, \end{aligned}$$

The initial values F_0 and V_0 are the current forward price and volatility, whereas W_1 and W_2 are two correlated Brownian motions with correlation coefficient ρ. The constant parameters β, α are such that $0 \leq \beta \leq 1$,$\alpha \geq 0$.

Once a particular stochastic volatility model is chosen, it must be calibrated against existing market data. Calibration is the process of identifying the set of model parameters which most likely give the observed data. For instance, in the Heston model, the set of model parameters $\{\omega, \theta, \eta, \rho\}$ can be estimated from historic underlying prices. Once the calibration has been performed, re-calibration of the model is needed over time.

2.1.3 Dupire's local volatility model

Given the computational complexity of stochastic volatility models and the difficulty in calibrating parameters to the market prices of vanilla options, people want to find a simpler way to price exotic options consistently. Since before Breeden and Litzenberger [7], it was well understood that the risk-neutral probability density function could be derived from the market prices of European options. Dupire [10] and Derman and Kani [9] show that under risk neutral measure, there exist a unique diffusion process consistent with these distributions.

In Dupire 's local volatility model, the constant volatility is replaced by a deterministic function of time and the stock price process, which is known as the local volatility function. The underlying price S_t is assumed to follow a stochastic differential equation

$$dS_t = rS_t\,dt + \sigma(t, S_t)S_t\,dW(t),$$

where r is the instantaneous risk free rate and $W(t)$ is a Brownian motion.

The diffusion coefficients $\sigma(t, S_t)$ are consistent with the market prices for all option prices on a given underlying. Dupire's formula shows

$$\sigma^2(K, T; S_0) = \frac{\frac{\partial C}{\partial T} + rK\frac{\partial C}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}}.$$

By construction, European style vanilla options whose values are determined by the marginal distributions can be priced correctly. In practice, this model is used to determine prices of exotic options which are consistent with observed prices of vanilla options.

In [11], Dupire showed that the local variance is a conditional expectation of instantaneous variance. Assume the stochastic process for stock prices is

$$dS_t = \mu_t S_t dt + \sqrt{\sigma_t} S_t dW(t),$$

he derived that

$$\sigma^2(K, T; S_0) = E[\sigma | S_T = K],$$

that is, local variance is the expectation of the instantaneous variance conditional on the final stock price S_T being equal to the strike price K. This equation implies that

Dupire's local volatility model is in fact a practical implementation of Gyöngy's theorem which we will introduce in the next section.

2.1.4 Volatility models with jumps

Diffusion based volatility models cannot explain why the implied volatility skew is so steep for very short expirations and why short-dated term structure of skew is inconsistent with any model. Therefore, jumps are necessary to be modeled. Examples of such models are Merton's jump diffusion model [21] and Bates' jump-diffusion stochastic volatility model [2]. In these models, the dynamics of the underlying is easy to understand and describe, since the distribution of jump sizes is known. They are easy to simulate using Monte Carlo method.

Merton's jump diffusion model

Assume the stock price follow the SDE

$$dS = \gamma S dt + \sigma S dW + (J-1) S dq,$$

where q is a Poisson process, $dq = 1$ with probability λdt, and $dq = 0$ with probability $1 - \lambda dt$. The jump size is lognormally distributed with mean log-jump μ and standard deviation δ. We can rewrite the SDE as

$$dS = \mu S dt + \sigma S dW + (e^{\alpha + \delta Z} - 1) S dq,$$

with $Z \sim N(0,1)$. This allows us to get the probability density of S_t

$$p_t(x) = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k \exp - \frac{(x - \gamma t - k\mu)^2}{2(\sigma^2 t + k\delta^2)}}{k! \sqrt{2\pi(\sigma^2 t + k\delta^2)}}.$$

Prices of European options in this model can be obtained as a series where each term involves a Black-Scholes formula.

Bates' model

Bates introduced the jump-diffusion stochastic volatility model by adding proportional log-normal jumps to the Heston stochastic volatility model. The model has the following form

$$dS_t = \mu_t S_t dt + \sqrt{V_t} S_t dW_1 + dJ_t$$
$$dV_t = \theta(\omega - V_t)dt + \eta\sqrt{V_t} dW_2,$$

$dW(t)^1$ and dW_2^t are Brownian motions with correlation ρ, and J_t is a compound Poisson process with intensity λ and log-normal distribution of jump sizes such that if k is its jump size then $\ln(1+k) \sim N(\ln(1+\bar{k}) - \frac{1}{2}\delta^2, \delta^2)$. The no-arbitrage condition fixes the drift of the risk neutral process, under the risk-neutral probability $\mu = r - \lambda\bar{k}$. Applying Itô's lemma, we obtain the equation for the log-price $X(t) = \ln S_t$,

$$dX(t) = (r - \lambda\bar{k} - \frac{1}{2}V_t)dt + \sqrt{V_t} dW(t)^1 + d\tilde{J}_t,$$

where $d\tilde{J}_t$ is a compound Poisson process with intensity λ and normal distribution of jump sizes.

Bates' model can also be viewed as a generalization of the Merton's jump diffusion model allowing for stochastic volatility. Although the no arbitrage condition fixes the drift of the price process, the risk-neutral measure is not unique, because other parameters of the model, for example, intensity of jumps and parameters of jump size distribution, can be changed.

2.2 Mimicking theorems

We want to construct simple processes which mimic certain features of the behavior of more complicated processes.

Since the European option prices are uniquely determined by the marginal distributions of the underlying price processes. In this section, we review some theorems in which the marginal distributions of a general process are matched by a Markov process.

2.2.1 Gyöngy's theorem

Dupire derived the local volatility formula using the forward equation. In an earlier work of Gyöngy [15], he developed a result that is considered as a rigorous proof of the existence of the local volatility model. He proved the following result.

Theorem 2.1. *(Gyöngy [15, Theorem 4.6]) Let W be an r-dimensional Brownian motion, and*

$$dX(t) = \mu_t dt + \sigma_t dW(t)$$

be a d-dimensional Itô process where μ is a bounded d-dimensional adapted process, and σ is a bounded $d \times r$-dimensional adapted process such that $\sigma\sigma^T$ is uniformly positive definite. There exist deterministic measurable function $\hat{\mu}$ and $\hat{\sigma}$ such that

$$\begin{aligned} \hat{\mu}(t, X(t)) &= E[\mu_t | X(t)] \text{ a.s. for each } t, \\ \hat{\sigma}\hat{\sigma}^T(t, X(t)) &= E[\sigma_t \sigma_t^T | X(t)] \text{ a.s. for each } t, \end{aligned}$$

and there exists a weak solution to the stochastic differential equation:

$$d\hat{X}_t = \hat{\mu}(t, \hat{X}_t)dt + \hat{\sigma}(t, \hat{X}_t)d\hat{W}(t)$$

such that $\mathscr{L}(\hat{X}_t) = \mathscr{L}(X(t))$ for all $t \in \mathbb{R}_+$, where $\mathscr{L}$ denotes the law of a random variable and $\hat{W}(t)$ denotes another Brownian motion, possibly on another space.

The assumptions on the drift and covariance processes in Gyöngy's theorem seem too restrictive for many applications. For example, Atlan [1] computed the conditional expectations for the Heston model using properties of the Bessel process and then used Gyöngy's theorem to ensure the existence of a diffusion with the same marginal distributions. However, the covariance process in the Heston model is a CIR process (2.1) which is neither bounded nor bounded away from zero, so the conditions of Gyöngy's theorem are not satisfied in this application.

Gyöngy's theorem is only valid for continuous Itô processes. It is important to extend the result to processes with jumps.

2.2.2 Brunick's theorem

In practice, the local volatility model is also used to price complex options with path dependent payoffs. However, the price cannot be uniquely determined by the marginal distributions of the asset price process. For example, the price of a barrier option would require knowledge of the joint probability distribution of the asset price process and its running maximum. Brunick [8] generalized Gyöngy's result under a weaker assumption so that the prices of path dependent options could be determined exactly.

Definition 2.1. *([8, Definition 2.1])*
Let $Y : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^n$ be a predictable process. We say that the path-functional Y is a ***measurably updatable statistic*** *if there exists a measurable function*

$$\phi : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^n$$

such that $Y(t+u, x) = \phi(t, Y(t,x); u, \Delta(t,x))$ for all $x \in C(\mathbb{R}_+, \mathbb{R}^d)$, where the map $\Delta : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^D) \to C(\mathbb{R}_+, \mathbb{R}^d)$ is defined by $\Delta(t,x)(u) = x(t,u) - x(t)$.

A measurably updatable statistic is a functional whose path-dependence can be summed up by a single vector in $\mathbb{R}^N$. $Y^1(t,x) = x(t)$ is an updatable statistic, as $Y^1(t+u,x) = Y^1(t,x) + \Delta(t,x)(u)$. Let $x^*(t) = \sup_{u \le t} x(u)$, we see that $Y^2(t,x) = [x(t), x^*(t)] \in \mathbb{R}^2$ is an updatable statistic as we can write

$$Y^2(t+u,x) = [x(t) + \Delta(t,x)(u), \max\{x^*(t), \sup_{v \in [0,u]} x(t) + \Delta(t,x)(v)\}].$$

Theorem 2.2. *([8, Theorem 2.11])*
Let W be an r-dimensional Brownian motion, and

$$dX(t) = \mu_t dt + \sigma_t dW(t)$$

be a d-dimensional Itô process where μ is a left-continuous d-dimensional adapted process, and σ is a left-continuous $d \times r$-dimensional adapted process with

$$E[\int_0^t |\mu_s| + |\sigma_s \sigma_s^T ds|ds] \le \infty \text{ for all } t.$$

Also suppose that $Y : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^n$ is a measurably updatable statistic such that the maps $x \mapsto Y(t,x)$ are continuous for each fixed t. Then there exist deterministic measurable function $\hat{\mu}$ and $\hat{\sigma}$ such that

$$\begin{aligned} \hat{\mu}(t, Y(t,X)) &= E[\mu_t | Y(t,X)] \text{ a.s. for Lebesgue-a.e. } t, \\ \hat{\sigma}\hat{\sigma}^T(t, Y(t,X)) &= E[\sigma_t \sigma_t^T | Y(t,X)] \text{ a.s. for Lebesgue-a.e. } t, \end{aligned}$$

and there exists a weak solution to the SDE:

$$d\hat{X}_t = \hat{\mu}(t, Y(t, \hat{X}))dt + \hat{\sigma}(t, Y(t, \hat{X}))d\hat{W}(t)$$

such that $\mathscr{L}(Y(t, \hat{X})) = \mathscr{L}(Y(t, X))$ *for all* $t \in \mathbb{R}_+$*, where* $\mathscr{L}$ *denotes the law of a random variable and* $\hat{W}(t)$ *denotes another Brownian motion.*

Brunick's theorem is more general than Gyöngy's. First, the requirements on μ and σ are weaker than the boundedness and uniform ellipticity in Gyögy's theorem. Secondly, this theorem implies the existence of a weak solution which preserves the one-dimensional marginal distribution of path-dependent functionals.

Corollary 2.1. *([8, Corallary 2.16]) Let* W *be an* r*-dimensional Brownian motion, and*

$$dX(t) = \mu_t dt + \sigma_t dW(t)$$

be a d*-dimensional Itô process where* μ *is a left-continuous* d*-dimensional adapted process, and* σ *is a left-continuous* $d \times r$*-dimensional adapted process with*

$$E[\int_0^t |\mu_s| + |\sigma_s \sigma_s^T ds|ds] \leq \infty \text{ for all } t.$$

Then there exist deterministic measurable function $\hat{\mu}$ *and* $\hat{\sigma}$ *such that*

$$\begin{aligned} \hat{\mu}(t, X(t)) &= E[\mu_t | X(t)] \text{ a.s. for Lebesgue-a.e. } t, \\ \hat{\sigma}\hat{\sigma}^T(t, X(t)) &= E[\sigma_t \sigma_t^T | X(t)] \text{ a.s. for Lebesgue-a.e. } t, \end{aligned}$$

and there exists a weak solution to the SDE:

$$d\hat{X}_t = \hat{\mu}(t, \hat{X}_t)dt + \hat{\sigma}(t, \hat{X}_t)d\hat{W}(t)$$

such that $\mathscr{L}(\hat{X}_t) = \mathscr{L}(X(t)))$ *for all* $t \in \mathbb{R}_+$*, where* $\mathscr{L}$ *denotes the law of a random variable and* $\hat{W}(t)$ *denotes another Brownian motion.*

This lemma relaxes the requirements on the coefficients of the Itô process in Gyöngy's theorem.

2.2.3 Bentata and Cont's theorems

Bentata and Cont [3], [4] extended Gyöngy's theorem to a discontinuous semimartingale. They showed that the flow of marginal distributions of a semimartingale X could be matched by the marginal distributions of a Markov process Y whose infinitesimal generator is expresses in terms of the local characteristics of X. They gave a construction of the Markov process Y as the solution to a martingale problem for an integral-differential operator. They applied these results to derive a partial integro-differential equation for call options in a general semimartingale model. This generalizes Dupire's local volatility formula.

Consider a semimartingale given by the decomposition

$$X(t) = X_0 + \int_0^t \beta_s^X ds + \int_0^t \sigma_s^X dW(s) + \int_0^t \int_{\|y\|\leq 1} y\widetilde{M}_X(dsdy) + \int_0^t \int_{\|y\|>1} yM_X(dsdy),$$

where W is a $\mathbb{R}^d$-valued Brownian motion, M_X is a positive, integer valued random measure on $[0,\infty) \times \mathbb{R}^n$ with compensator μ_X , $\widetilde{M}_x = M_X - \mu_X$ is the compensated measure, β_t^X and σ_t^X are adapted processes in $\mathbb{R}^n$ and $M_{n\times d}(\mathbb{R})$.

Define, for $t \geq 0$, $z \in \mathbb{R}^d$

$$a^Y(t,z) = E[\sigma_t^X(\sigma_t^X)^T | X(t-) = z]$$

$$b(x,t) = E[\beta(t)|X(t-) = z]$$

$$m_Y(t,y,z) = E[m_X(t,y)|X(t-) = z].$$

Let M_Y be an integer valued random measure on $[0,T] \times \mathbb{R}^d$ with compensator $m_Y(t,y,Y(t-))dydt$, $\tilde{M}_Y = M_Y - m_y$ the associated compensated random measure, $\sigma^Y : [0,\infty) \times \mathbb{R}^d \mapsto M_{d\times n}(\mathbb{R})$ is a measurable function such that $\sigma^Y(t,z)(\sigma^Y(t,z))^T = a^Y(t,z)$.

Theorem 2.3. *([3, Theorem 1])*

If the function β^Y, a^Y and m_Y are continuous in (t,z) on $[0,T] \times \mathbb{R}^d$, the stochastic *differential equation*

$$\begin{aligned} Y(t) =& X_0 + \int_0^t \beta_s^Y(s, Y(s))ds + \int_0^t \sigma_s^Y(s, Y(s))dW(s) + \int_0^t \int_{\|y\|\leq 1} y\widetilde{M}_Y(dsdy) \\ &+ \int_0^t \int_{\|y\|>1} yM_Y(dsdy), \end{aligned}$$

admits a weak solution $(Y(t))_{t\in[0,T]}$ *whose marginal distributions mimic those of* X*:*

$$Y(t) = X(t) \quad \text{in distribution for any} \ \ t \in [0,T].$$

In the proof of this theorem, Bentata and Cont showed that, for any C^2 function f with compact support

$$\begin{aligned} E[f(X(t))] =& f(X_0) + \int_0^T E[\nabla f(X(t-)) \cdot \beta^Y(t, X(t-))]dt \\ &+ \frac{1}{2}\int_0^T E[\mathrm{tr}[\nabla^2 f(X(t-))a^Y(t, X(t-))]]dt \\ &+ \int_0^T \int_{\mathbb{R}^d} E[(f(X(t-) + y) - f(X(t-)) \\ &- 1_{\{\|y\|\leq 1\}} y\nabla f(X(t-)))m_Y(t, dy, X(t-))]dt. \end{aligned}$$

And it is easy to see that for the mimicking $Y(t)$, we also have

$$\begin{aligned} E[f(Y(t))] =& f(X_0) + \int_0^T E[\nabla f(Y(t-)) \cdot \beta^Y(t, Y(t-))]dt \\ &+ \frac{1}{2}\int_0^T E[\mathrm{tr}[\nabla^2 f(Y(t-))a^Y(t, Y(t-))]]dt \\ &+ \int_0^T \int_{\mathbb{R}^d} E[(f(Y(t-) + y) - f(Y(t-)) \\ &- 1_{\{\|y\|\leq 1\}} y\nabla f(Y(t-)))m_Y(t, dy, Y(t-))]dt. \end{aligned}$$

In order to show that $X(t)$ and $Y(t)$ have the same marginal distributions, they asserted that

$$E[f(X(t))] = E[f(Y(t))],$$

for any $f \in C_0^2$. However, they did not provide the proof of this statement in [3].

In [3], Y is constructed as the solution $((Y(t))_{t\in[0,T]}, Q_{X_0})$ of the martingale problem for an integro-differential operator

$$Lf(t,x) = \sum_{i=1}^{d} \beta_i^Y(t,x)\frac{\partial f}{\partial x_i}(t,x) + \sum_{i,j=1}^{d} \frac{a_{ij}^Y(t,x)^2}{2}\frac{\partial^2 f}{\partial x_i \partial x_j}(t,x)$$
$$+ \int_{\mathbb{R}^d} [f(t,x+y) - f(t,x) - \sum_{i=1}^{d} 1_{\{\|y\|\le 1\}} y_i \frac{\partial f}{\partial x_i}(t,x)] m_Y(t,y,x)dy.$$

For any $f \in E \subset \mathrm{dom}(L)$,

$$M_t^f = f(Y(t)) - f(x) - \int_0^t Lf(X(u))du$$

is a Q_x martingale.

The uniqueness if $Y(t)$ is guaranteed by the following theorem.

Theorem 2.4. *([27]) If either*

(i) for any $t \in [0,T]$, $x, z \in \mathbb{R}^d$, $x^T a_t^X x > 0$,

(ii) for any $\beta > 0$, $c > 0$, $m(t,y,\omega) \ge \frac{c}{|y|^{1+\beta}}$,

then Y is the unique Markov process with infinitesimal generator L.

Bentata and Cont also derived a Dupire type formula for an asset price S whose dynamics under the pricing measure is a stochastic volatility model with random jumps,

$$dS_T = S_0 \int_0^T r(t)S_{t-}dt + \int_0^T S_{t-}\delta_t dW(t) + \int_0^T \int_{-\infty}^{\infty} S_{t-}(e^y - 1)\tilde{M}(dtdy),$$

where $r(t)$ is the discount rate, δ_t the spot volatility process and $\tilde{M}$ is a compensated random measure with compensator

$$\mu(\omega; dtdy) = m(\omega; t, y)dtdy.$$

Theorem 2.5. *([4, Propostion 3])*

Assume $r(t)$ and δ_t are locally bounded process and that for any $T > 0$, there exists a constant $0 < c_T < \infty$ such that $\int_0^T \int_{y>1} e^{2y} m(\cdot; t, y)dy \le c_T$, $\int (1 \wedge y^2) m_X(\cdot; t, y)dy \le c_T$. Define

$$\sigma(t,z) = \sqrt{E[\delta_t^2 | S_{t-} = z]},$$
$$\nu(t,y,z) = E[m(t,y)|S_{t-} = z].$$

The value $C_t(T,K)$ *at time* t *of a call option with expiry* $T > t$ *and strike* $K > 0$ *is given by*

$$C_t(T,K) = E[\max(S_T - K, 0)|\mathcal{F}_t].$$

Then the call option price $(T,K) \mapsto C_t(T,K)$ *as a function of maturity and strike, is a solution (in the sense of distributions) of the partial integro differential equation on* $[t,\infty) \times (0,\infty)$,

$$\begin{aligned}\frac{\partial C}{\partial T} =& \frac{K^2\sigma(T,K)^2}{2}\frac{\partial^2 C}{\partial K^2} - rK\frac{\partial C}{\partial K} \\ &+ \int_{-\infty}^{\infty} dyn(T,y,K)e^y\Big\{C(T,Ke^{-y}) - C(T,K) + K(1-e^{-y})\frac{\partial C}{\partial K}\Big\}\end{aligned}$$

with initial condition $C(t,K) = (S_t - K)^+$ *for any* $K > 0$.

This partial integro-differential equation generalizes the Dupire formula derived in [10] for continuous process to the case of semimartingales with jumps. It implies that, any arbitrage free option price across strike and maturity may be parameterized by a local volatility function $\sigma(t,S)$ and a local Levy measure $m(t,S,z)$.

They also gave some examples of stochastic models, including marked point processes and time changed Levy processes.

In my thesis, we consider the same mimicking theorem of semimartingales as Bentata and Cont did. However, my approach is totally different. In [3], Bentata and Cont concluded that $E[f(X(t))] = E[f(Y(t))]$ for any C^2 function f, therefore, $X(t)$ and $Y(t)$ have the same marginal distributions. In my proof, I show that $p_X(x,t)$ and $p_Y(x,t)$, the probability density functions of $X(t)$ and $Y(t)$, satisfies the same partial integro-differential equation. And then I prove that this equation has a unique fundamental solution under some assumptions for the diffusion and variance coefficients.

In [25], Ming Shi extended Gyögy's theorem to pure jump processes and applied his results to multi-name credit modeling. He also talked about the mimicking theorem of semimartingales in Section 3.4 which was a joint work with me. We started with the function $f(x) = e^{-ix\xi}$ for $x, \xi \in \mathbb{R}^n$. $E[f(X(t))]$ is the Fourier transform of the marginal distribution function $p_X(x,t)$. We then used this fact to derive the forward equation satisfied by $p_X(x,t)$.

In my thesis, I generalize this result by choosing f as any C^2 function with compact support.

Chapter 3

A Proof of Gyöngy's Theorem

Gyögy proved his theorem by extending a result of Krylov [20]. We summarize his main ideas of proof here. Consider the Green measure of an Itô process $X(t)$ with killing rate $\gamma(t)$, which is defined by

$$\mu(A) = E\Big[\int_0^{+\infty} 1_A(X(t))e^{-\int_0^t \gamma(s,\omega)ds}dt\Big],$$

for every Borel set $A \subset R^n$, where 1_A denotes the indicator function of the set A, γ is a non-negative $\mathfrak{t}$-adapted stochastic process. Denote the mimicking process by $Y(t)$, Gyöngy showed that the Green measure of $(t, X(t))$ is identical the the Green measure of $(t, Y(t))$. Let $\gamma(t) \equiv 1$, then we have

$$E\left[\int_0^{+\infty} e^{-t} f(t, X(t))dt\right] = E\left[\int_0^{+\infty} e^{-t} f(t, Y(t))dt\right]$$

for every bounded non-negative Borel measurable function f. Taking $f(t, x) = e^{-\lambda t}g(x)$ with arbitrary non-negative constant λ and functions $g \in C_0(\mathbb{R}^n)$, we get

$$\int_0^{+\infty} e^{-\lambda t}e^{-t}E[g(X(t))]dt = \int_0^{+\infty} e^{-\lambda t}e^{-t}E[g(Y(t))]dt,$$

this gives us

$$E[g(X(t))] = E[g(Y(t))].$$

Hence it follows that the distribution of $X(t)$ and $Y(t)$ are same for every t. However, the proof that $X(t)$ and $Y(t)$ have the same Green measure is quite long and technical.

We present an intuitive proof of Gyöngy's theorem. The proof is based on the uniqueness of solutions to a parabolic equation. Assume $X(t)$ has probability density functions $p_X(x, t)$. We first derive the partial differential equation satisfied by the marginal distributions of an Itô process. Then we construct a Markov process and show that the marginal distributions of this Markov process satisfy the same partial differential equation. By uniqueness of solutions to a parabolic equation [14], we conclude that the Markov process has the same marginal distributions as the Itô process.

3.1 Uniqueness of a solution to a partial differential equation

In this section, we give a proof of Gyöngy's theorem using uniqueness of solutions to a partial differential equation. We shall use the following definition and theorems.

Definition 3.1. *(Fundamental solutions of a parabolic equation)[14, Sections 1.1 & 1.6] Suppose a_{ij}, b_i and c are $\mathbb{R}$-valued functions on Ω. A fundamental solution of*

$$Lu := \sum_{i,j}^{n} a_{ij}(x,t)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x,t)\frac{\partial u}{\partial x_i} + c(x,t)u - \frac{\partial u}{\partial t} = 0 \tag{3.1}$$

in $\Omega = \bar{D} \times [T_0, T_1]$ is a $\mathbb{R}$-valued function $\Gamma(x,t;y,\tau)$ defined for all $(x,t) \in U$, $(y,\tau) \in U$, $t > \tau$, which satisfies the following conditions:
(i) for fixed $(y,\tau) \in \Omega$, it satisfies (3.1), as a function of (x,t), $x \in D$, $\tau < t < T$;
(ii) for every continuous $\mathbb{R}$-valued function $f(x)$ in $\bar{D}$ such that [14, Section 1.6, Equation (6.1)],

$$|f(x)| \leq const. \exp(h|x|^2),$$

if $x \in D$, for some positive constant h. Then

$$\lim_{t \to \tau} \int_D \Gamma(x,t;y,\tau) f(y) dy = f(x).$$

The integrals in Definition 3.1 exist only if $4h(t-\tau) < \lambda$, where $\lambda = \lambda_0$ is defined as in [14, Section 1.2, Equation (2.2)] and the proof of Lemma 4.2 in this thesis. If

$$h < \frac{\lambda_0}{4(T_1 - T_0)}$$

then the integrals in Definition 3.1 exist for all $T_0 \leq \tau < t \leq T_1$ [14, Section 1.6].

Definition 3.2. *[14] A function $f(x,t)$ is said to be Hölder continuous with exponent α if*

$$|f(x,t) - f(x_0,t_0)| \leq C(|x - x_0|^\alpha + |t - t_0|^{\alpha/2})$$

for any $(x,t), (x_0,t_0) \in [T_0,T_1] \times \mathbb{R}^n$ some constant $C > 0$.

Theorem 3.1. *[14, Theorem 1.15]*

Assume that L is parabolic and the coefficients of L

$$a_{ij}, \frac{\partial a_{ij}}{\partial x_k}, \frac{\partial^2 a_{ij}}{\partial x_k \partial x_l}, b_i, \frac{\partial b_i}{\partial x_k}, c$$

are bounded, continuous $\mathbb{R}$-valued functions on $\mathbb{R}^n \times [T_0, T_1]$; they satisfy a uniform Hölder condition with exponent α in $x \in \mathbb{R}^n$. a_{ij} satisfies the ellipticity condition.

Then a fundamental solution Γ^* of $L^* u = 0$ exists and

$$\Gamma(x, t; y, \tau) = \Gamma^*(y, \tau; x, t),$$

where L^* is the formal adjoint of L,

$$L^* u = \sum_{i,j}^{n} \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}(x,t)u) + \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(b_i(x,t)u) + c(x,t)u + \frac{\partial u}{\partial t}$$

Theorem 3.2. *([13, Theorem 3.4]) The fundamental solution of (3.1) is unique.*

Proof. Suppose that $\Gamma(x, t; y, \tau)$, $\tilde{\Gamma}(x, t; y, \tau)$ are two fundamental solutions for $Lu = 0$. Applying [14, Theorem 1.15], we see that

$$\Gamma(x, t; y, \tau) = \Gamma^*(y, \tau; x, t) = \tilde{\Gamma}(x, t; y, \tau)$$

for any $(x, t), (y, \tau) \in [T_0, T_1] \times \mathbb{R}^n$, $t > \tau$.

And so $\Gamma(x, t; y, \tau) = \tilde{\Gamma}(x, t; y, \tau)$ as desired. □

Now we can proceed to give another proof of Gyöngy's theorem.

Theorem 3.3. *Let W be an r-dimensional Brownian motion on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and*

$$dX(t) = \mu_t dt + \sigma_t dW(t)$$

be an $\mathbb{R}^n$-valued Itô process, μ is a bounded $\mathbb{R}^n$-valued adapted process, and σ is a bounded $\mathbb{R}^{n \times d}$-valued adapted process such that $\sigma\sigma^T$ is uniformly positive definite. Then there exist deterministic measurable function $a(t, x)$ and $b(t, x)$ such that

$$\begin{aligned} b(t,x) &= E[\mu_t | X(t) = x] \text{ a.s. for } t \geq 0, \\ a(t,x)a^T(t,x) &= E[\sigma_t \sigma_t^T | X(t) = x] \text{ a.s. for } t, \end{aligned}$$

and

$$a_{ij}, \frac{\partial a_{ij}}{\partial x_k}, \frac{\partial^2 a_{ij}}{\partial x_k \partial x_l}, b_i, \frac{\partial b_i}{\partial x_k}$$

are bounded and Hölder continuous with exponent α*,* $0 < \alpha < 1$*. Assume* $X(t)$ *has probability density functions* $p_X(x,t)$*. Then there exists a weak solution to the stochastic differential equation:*

$$dY(t) = b(t, Y(t))dt + a(t, Y(t))d\hat{W}(t)$$

such that $\mathscr{L}(Y(t)) = \mathscr{L}(X(t))$ *for all* $t > 0$*, where* $\mathscr{L}$ *denotes the law of a random variable and* $\hat{W}(t)$ *denotes another Brownian motion, on another probability space* $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$.

Proof. For any function $f \in C_0^2(R^n)$, Itò formula shows

$$\begin{aligned} f(X_T) - f(x) &= \int_0^T \sum_{i=1}^n \frac{\partial}{\partial x_i} f(X_t) dX(t) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n \int_0^T \frac{\partial^2}{\partial x_i \partial x_j} f(X_t) d[X^i, X^j]_t \\ &= \int_0^T \sum_{i=1}^n \frac{\partial}{\partial x_i} f(X_t)(\mu_t dt + \sigma_t dW(t)) \\ &\quad + \frac{1}{2} \sum_{i,j=1}^n \int_0^T \frac{\partial^2}{\partial x_i \partial x_j} f(X_t)(\sigma(t)\sigma^T(t))_{ij} dt \end{aligned}$$

Taking expectations on both side, we get

$$\begin{aligned} E[f(X_T)] &= f(x) + E\Big[\int_0^T \sum_{i=1}^n \frac{\partial}{\partial x_i} f(X_t)\mu_t dt\Big] \\ &\quad + \frac{1}{2} E\Big[\int_0^T \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} f(X_t)(\sigma(t)\sigma^T(t))_{ij} dt\Big], \\ &= f(x) + \int_0^T E\Big[\sum_{i=1}^n \frac{\partial}{\partial x_i} f(X_t)\mu_t\Big] dt \\ &\quad + \frac{1}{2} \int_0^T \sum_{i,j=1}^n E\Big[\frac{\partial^2}{\partial x_i \partial x_j} f(X(t-))(\sigma(t)\sigma^T(t))_{jk}\Big] dt]. \end{aligned}$$

Using iterated conditional expectation and conditioning on X_t gives

$$\begin{aligned}E[f(X_T)] =& f(x) + \int_0^T E\Big[\sum_{i=1}^n \frac{\partial}{\partial x_i} f(X_t) E[\mu_t|X_t]\Big] dt \\ &+ \frac{1}{2}\int_0^T \sum_{j,k=1}^n E\Big[\frac{\partial^2}{\partial x_i \partial x_j} f(X_t) E[(\sigma(t)\sigma^T(t))_{ij}|X_t]]dt\Big] \\ =& f(x) + \int_0^T E\Big[\sum_{i=1}^n \frac{\partial}{\partial x_i} f(X_t) b(X_t,t) + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} f(X_t) a_{ij}(X_t,t)\Big] dt.\end{aligned}$$

Denoting

$$Af(x) = \sum_{i=1}^n b_i(x,t)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x,t)\frac{\partial^2 f}{\partial x_i x_j},$$

we have

$$E[f(X_T)] = f(x) + \int_0^T E\Big[Af(X(t))\Big] dt. \tag{3.2}$$

If $X(t)$ has probability density functions $p_X(x,t)$, we can rewrite (3.2) as

$$\int_{\mathbb{R}^n} f(x) p_X(x,T) dx = f(x) + \int_0^T \int_{\mathbb{R}^n} Af(x) p_X(x,t) dx dt. \tag{3.3}$$

Taking derivatives of (3.3) with respect to T, we get

$$\begin{aligned}\frac{\partial}{\partial T}\int_{\mathbb{R}^n} f(x) p_X(x,T) dx &= \int_{\mathbb{R}^n} \big(Af(x)\big) p_X(x,T) dx \\ &= \int_{\mathbb{R}^n} \Big(\sum_{i=1}^n b_i(x,T)\frac{\partial f}{\partial x_i} p_X(x,T) \\ &+ \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x,T)\frac{\partial^2 f}{\partial x_i x_j} p_X(x,T)\Big) dx.\end{aligned}$$

Integrating by parts on the right hand side gives

$$\begin{aligned}&\frac{\partial}{\partial T}\int_{\mathbb{R}^n} f(x) p_X(x,T) dx \\ &= \int_{\mathbb{R}^n} \Big(-\sum_{i=1}^n f(x)\frac{\partial}{\partial x_i}(b(x,T)p_X(x,T)) + \frac{1}{2}\sum_{i,j=1}^n f(x)\frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}(x,T)p_X(x,T))\Big) dx \\ &= \int_{\mathbb{R}^n} f(x)\Big(-\sum_{i=1}^n \frac{\partial}{\partial x_i}(b(x,T)p_X(x,T)) + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}(x,T)p_X(x,T))\Big) dx \\ &= \int_{\mathbb{R}^n} f(x)(A^* p_X(x,T)) dx\end{aligned} \tag{3.4}$$

where A^* is the formal adjoint of A,

$$A^* p_X(x,t) := -\sum_{i=1}^n \frac{\partial}{\partial x_i}(b(x,t)p_X(x,t)) + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}(x,t)p_X(x,t)).$$

Since (3.4) is true for any C_0^2 function f, then $p_X(x,t)$ is a weak solution of

$$\frac{\partial p_X}{\partial t} = A^* p_X(x,t).$$

Consider the process $Y(t)$ defined as the solution to

$$dY(t) = b(t, Y(t))dt + a(t, Y(t))d\hat{W}(t).$$

Itô formula gives

$$E[f(Y_T)] = f(x) + \int_0^T E\Big[\sum_{i=1}^n \frac{\partial}{\partial x_i} f(Y_t) b(Y_t, t) + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} f(Y_t) a_{ij}(Y_t, t)\Big] dt.$$

Let $p_Y(x,t)$ denotes the probability density functions of Y_t, we have

$$\int_{\mathbb{R}^n} f(x) p_Y(x,T)dx = f(x) + \int_0^T \int_{\mathbb{R}^n} Af(x) p_Y(x,t) dx dt. \tag{3.5}$$

Then we take derivatives with respect to T and integrate by parts, we obtain

$$\frac{\partial}{\partial T}\int_{\mathbb{R}^n} f(x) p_Y(x,T)dx = \int_{\mathbb{R}^n} f(x)(A^* p_Y(x,T))dx.$$

Therefore, $p_Y(x,t)$ is a solution to

$$\frac{\partial p_Y}{\partial t} = A^* p_Y(x,t),$$

in the weak sense, with initial condition $\tilde{p}(x,0) = \delta(x)$, where $\delta(x)$ is the Dirac delta function.

We now consider an initial value problem

$$\begin{aligned} \frac{\partial u}{\partial t} &= A^* u(x,t) \\ u(x,0) &= \delta(x). \end{aligned} \tag{3.6}$$

Since $p_X(x,t)$, $p_Y(x,t)$ are solutions of (3.6), it is enough to show that (3.6) has a unique fundamental solution.

In [14, Section 1.4], a fundamental solution of

$$Lu := A^* u - \frac{\partial u}{\partial t},$$

is constructed by the parametrix method. By Theorem 3.2, we obtain the uniqueness of the fundamental solution. □

Chapter 4

Forward Equation for a Semimartingale

In this chapter, we develop the main theorem of this dissertation. We first introduce the definition of semimartingales and define the Markovian projection of semimartingales. Then we derive the generator of a semimartingale and the backward equation. When the semimartingale has a probability density function, we show that this function satisfies the forward equation, the adjoint of the backward equation. Next we construct a fundamental solution of the forward equation, through our construction, discover the conditions which ensure that this equation has fundamental solutions. These conditions guarantee the existence of the probability density functions for the semimartingale. Finally, we show that the fundamental solution of the forward equation is unique. Our existence and uniqueness result implies that the mimicking process has the same marginal distributions as the original semimartingale.

4.1 Semimartingales and Markovian projection

In this section, we give the definition of a semimartingale and define its Markovian projection.

Definition 4.1. *(Semimartingale, [24, Definition IV.15.1])*
A process X is a semimartingale on $(\Omega, \mathcal{F}, (\mathbb{F}(t))_{t\geq 0}, \mathbb{P})$ if it can be written in the form:

$$X(t) = X_0 + M_t + A_t, \tag{4.1}$$

where M is a local martingale null at 0 with càdlàg paths and A is an adapted process with paths of finite variation, and the filtration is right-continuous.

An $\mathbb{R}^n$-valued process $X = (X^1, \ldots, X^n)$ is a semimartingale if each of its components X^i is a semimartingale.

We emphasize that the decomposition (4.1) need not be unique. Semimartingales are good integrators, forming the largest class of processes with respect to which the Itô integral can be defined. The class of semimartingales is quite large, including, for example, all Itô processes and Lévy processes.

Proposition 4.1. *(Itô decomposition for semimartingales, [4, Equation (2)])* *On a filtered probability space* $(\Omega, \mathcal{F}, (\mathbb{F}_t)_{t\geq 0}, \mathbb{P})$*, a semimartingale can be given by the decomposition*

$$X(t) = X(0) + \int_0^t \beta(s)ds + \int_0^t \sigma(s)dW(s) + \int_0^t \int_{\|y\|\leq 1} y\widetilde{M}(dsdy) + \int_0^t \int_{\|y\|>1} yM(dsdy), \tag{4.2}$$

where W *is a* $\mathbb{R}^d$*-valued Brownian motion,* M *is a positive, integer valued random measure on* $[0,\infty)\times\mathbb{R}^n$ *with compensator* μ *,* $\widetilde{M} = M - \mu$ *is the compensated measure,* μ *has a density* $m(t,\omega,y)dy$*,* $\beta(t)$ *and* $\sigma(t)$ *are adapted processes valued in* $\mathbb{R}^n$ *and* $M_{n\times d}(\mathbb{R})$*.*

Proposition 4.2. *(Itô formula for semimartingales, [23, Theorem 71])* *Let* $(X(t))_{t\geq 0}$ *be a semimartingale. For any twice continuously differentiable function* $f : \mathbb{R}^n \to \mathbb{R}^n$*,*

$$\begin{aligned} f(X_T) - f(X(0)) = &\sum_{i=1}^n \int_0^T \frac{\partial}{\partial x_i} f(X(t-))dX(t)^i + \frac{1}{2}\sum_{i,j=1}^n \int_0^T \frac{\partial^2}{\partial x_i \partial x_j} f(X(t-))d[X^i, X^j]_t \\ &+ \sum_{0\leq t\leq T}\left[f(X(t-) + \triangle X(t)) - f(X(t-)) - \sum_{i=1}^n \frac{\partial}{\partial x_i} f(X(t-))\triangle X(t)^i \right] \end{aligned}$$

We want to construct a stochastic differential equation whose solution mimics the marginal distributions of a semimartingale. We call the mimicking process a *Markovian projection of a semimartingale.*

Gyöngy [15] showed that there exists a Markovian projection of an Itô process. Brunick [8] generalized Gyöngy's result under a weaker assumption, and proved that there exists a Markovian projection of a two-dimensional process. Bentata and Cont [3], [4] gave the existence of the Markovian projection of a semimartingale. The Markovian projection is constructed as a solution to a martingale problem.

Here we present a partial differential equation based proof of mimicing theorem. Consider a semimartingale $X(t)$ with decomposition (4.2) and further require

(A1) $\sigma(t)\sigma(t)^*$ satisfies the uniformly ellipticity condition for any $0 \leq t \leq T$.

That is, there is a constant $\lambda > 0$ such that

$$\sum_{i,j=1}^{n} (\sigma(t)\sigma(t)^T)_{ij}\xi_i\xi_j \geq \lambda|\xi|^2,$$

for all $\xi \in \mathbb{R}^n$ and any $0 \leq t \leq T$. Ellipticity thus means that the symmetric matrix $\sigma(t)\sigma(t)^* \in \mathbb{R}^{n\times n}$ is positive definite, with smallest eigenvalue greater than or equal to $\lambda > 0$.

(A2) For any $T > 0$, $\beta(t)$ and $\sigma(t)$ are bounded functions of $t \in [0, T]$.

Suppose the process $X(t)$ start at $X(t_0) = x_0$, and the density function of $X(t)$ is $p_X(t_0, x_0; t, \cdot)$, we denote the density by $p_X(t_0, x_0; t, x)$ and abbreviate by $p_X(t, x)$

Theorem 4.1. *(Markovian projection of a semimartingale)*
Let $X(t)$ be a semimartingale starting with $X(t_0) = x_0$, and $X(t)$ has a decomposition

$$X(t) = X(t_0)+\int_{t_0}^{t} \beta(s)ds+\int_{t_0}^{t} \sigma(s)dW(s)+\int_{t_0}^{t}\int_{\|y\|\leq 1} y\widetilde{M}(dsdy)+\int_{t_0}^{t}\int_{\|y\|>1} yM(dsdy),$$

where $\beta(t)$, $\sigma(t)$ satisfy (A1) and (A2). Assume $X(t)$ has probability density functions $p_X(t, x)$.

Define

$$a(x, t) := E[\sigma(t)\sigma^*(t)|X(t-) = x]$$

$$b(x, t) := E[\beta(t)|X(t-) = x]$$

$$n(t, A, x) := E[m(t, A, \cdot)|X(t-) = x],$$

for all $t \geq 0$, $x \in \mathbb{R}^n$, and any Borel set $A \subset \mathbb{R}^n$. Let N be a positive, integer-valued random measure on $[0, \infty) \times \mathbb{R}^n$ with compensator n, where n has a density $\nu(t, \omega, x)dx$, and $\widetilde{N} = N - \nu$ is the associated compensated random measure.

We assume that

(i) a, b, and $A \to n(t, A, x)$ are continuous in (t, x) on $[0, \infty) \times R^n$,

(ii) there is a constant K, such that $\int_R^n m(t, \cdot, dy) \leq K < \infty$ a.s.

(iii) there is no jump exactly at t for any fixed t with probability 1.

(iv) ν *has a compact support.*

Define a process $Y(t)$

$$Y(t) = X(t_0) + \int_{t_0}^{t} b(s, Y(s))ds + \int_{t_0}^{t} a^{\frac{1}{2}}(s, Y(s))d\widetilde{W}_s + \int_{t_0}^{t} \int_{\|y\| \le 1} y\widetilde{N}(dsdy)$$
$$+ \int_{t_0}^{t} \int_{\|y\|>1} yN(dsdy)$$

where $\widetilde{W}_t$ *is an* $\mathbb{R}^n$*-valued Brownian motion. Then* $Y(t)$ *is a Markovian projection of the semimartingale* $X(t)$*,* $Y(t)$ *and* $X(t)$ *have the same marginal distributions for any* $0 \le t \le T$*.*

4.2 Forward equation

Suppose $X(t)$ is a semimartingale with generator A, $Y(t)$ is the process defined in Theorem 4.1 with generator $\hat{A}$. We want to show that the density $p_X(t_0, x_0; t, x)$ of $X(t)$ and $p_Y(t_0, x_0; t, y)$ are solutions to the forward equation defined by $\hat{A}$.

Definition 4.2. *The generator* $\hat{A}_t, t \ge 0$ *of a (time-inhomogeneous) process* $(Y(t))_{t \ge 0}$ *is defined by*

$$\hat{A}_{t_0} f(x) := \lim_{t \to t_0} \frac{E^x[f(Y(t))] - f(x)}{t}, \qquad x \in \mathbb{R}^n, t_0 \ge 0,$$

where E^x *is the expectation with respect to the probability law for* $Y(t)$ *starting at* $Y(t_0) = x$*. The set of functions* $f : \mathbb{R}^n \to \mathbb{R}$ *such that the limit exists at* (t_0, x) *is denoted by* $\mathscr{D}_{\hat{A}_{t_0}}(x)$*, while* $\mathscr{D}_{\hat{A}}$ *is the set of functions for which the limit exists for all* $x \in \mathbb{R}^n$ *and* $t_0 \ge 0$*.*

We assume $\hat{A}$ has the property that $\mathscr{D}_{\hat{A}} \subset C_\infty(R^n, R)$ is dense, where $C_\infty(R^n, R)$ is the set of continuous functions from R^n to R which vanish at infinity.

If $Y(t)$ also obeys the conditions

(i) $Y(t)$ is time-homogeneous;

(ii) $(\hat{A}, \mathscr{D}_{\hat{A}})$ satisfies the positive maximum principle;

(iii) $R(\lambda - \hat{A})$ is dense in $C_\infty(R^n, R)$ for some $\lambda > 0$;

then by [18, Theorem 4.5.3], $\hat{A}$ is a generator of a Feller semigroup and $Y(t)$ is a time-homogeneous Feller process.

The Courrège theorem [18, Theorem 4.5.21] shows that if $A : C_\infty(R^n, R) \to C(R^n, R)$ is a linear operator satisfying the positive maximum principle. Then there exist functions $a, b, c : R^n \to R$ and a kernel μ such that for $u \in C_0^\infty(R^n, R)$

$$Au(x) = \sum_{k,l=1}^{n} a_{kl}(x)\frac{\partial^2 u(x)}{\partial x_k \partial x_l} + \sum_{j=1}^{n} b_j(x)\frac{\partial u(x)}{\partial x_j} + c(x)u(x)$$
$$+ \int_{R^n} \{u(y) - \chi(y-x)u(x) - \sum_{j=1}^{n} \frac{\partial u(x)}{\partial x_j}\chi(y-x)(y_j - x_j) \times \mu(x, dy),$$

where $\chi \in C_\infty(R^n, R)$ with $0 \le \chi \le 1$ and $\chi|_{B_1(0)} = 1$.

The Courrège theorem gives us an example of the structure of $\hat{A}$. Conversely, when $\hat{A}$ has the structure described in the Courrège theorem, then $\hat{A}$ is the generator of a time-homogeneous Feller process.

Proposition 4.3. *Let X be a semimartingale with decomposition (4.2) and satisfy the assumptions in Theorem 4.1. If $f \in C_0^2(\mathbb{R}^n)$, that is f is a C^2 function with compact support on $\mathbb{R}^n$, then $f \in \mathscr{D}_{\hat{A}}$ and the generator of $Y(t)$ is*

$$\hat{A}_t f(x) = \sum_{i=1}^{n} b_i(x,t)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x,t)\frac{\partial^2 f}{\partial x_i x_j}$$
$$+ \int_{\mathbb{R}^n} \big(f(x+y) - f(x) - 1_{\|y\|\le 1}\sum_{i=1}^{n} y_i \frac{\partial f}{\partial x_i}\big)\nu(t,y,x)dy.$$

Proof. Since X is a semimartingale, with $X(t-) = x$, the Itô formula gives

$$f(X_T) - f(x)$$
$$= \int_t^T \nabla f(X(s-)) \cdot dX(s) + \frac{1}{2}\sum_{i,j=1}^{n}\int_t^T \frac{\partial^2}{\partial x_i \partial x_j} f(X(s-))d[X^i, X^j]_s$$
$$+ \sum_{0\le s\le T} [f(X(s-) + \triangle X(s)) - f(X(s-)) - \sum_{i=1}^{n}\frac{\partial}{\partial x_i} f(X(s-))\triangle X(s)^i]$$
$$= \int_t^T \nabla f(X(s-)) \cdot \left[\beta(s)ds + \sigma(s)dW(s) + \int_{\|y\|\le 1} y\widetilde{M}(dsdy) + \int_{\|y\|>1} yM(dsdy)\right]$$
$$+ \frac{1}{2}\sum_{i,j=1}^{n}\int_t^T \frac{\partial^2}{\partial x_i \partial x_j} f(X(s-))d[X^i, X^j]_s$$
$$+ \sum_{0\le t\le T}\left[f(X(s-) + \triangle X(s)) - f(X(s-)) - \sum_{i=1}^{n}\frac{\partial}{\partial x_i} f(X(s-))\triangle X(s)^i\right].$$

Simplifying, we obtain

$$
\begin{aligned}
&f(X_T) - f(x) \\
=&\int_t^T \nabla f(X(s-)) \cdot \beta(s) ds + \int_t^T \nabla f(X(s-)) \cdot \sigma(s) dW(s) \\
&+ \frac{1}{2} \int_t^T \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} f(X(s-))(\sigma(s)\sigma^T(s))_{ij} ds + \int_t^T \int_{\|y\| \leq 1} y \cdot \nabla f(X(s-)) \widetilde{M}(dsdy) \\
&+ \int_t^T \int_{\|y\|>1} y \cdot \nabla f(X(s-)) M(dsdy) \\
&+ \int_t^T \int_{\mathbb{R}^n} (f(X(s-)+y) - f(X(s-)) - y \cdot \nabla f(X(s-))) M(dsdy).
\end{aligned}
$$

We add the last two terms in the preceding equation and get

$$
\begin{aligned}
&f(X_T) - f(x) \\
=&\int_t^T \nabla f(X(s-)) \cdot \beta(s) ds + \int_t^T \nabla f(X(s-)) \cdot \sigma(s) dW(s) \\
&+ \frac{1}{2} \int_t^T \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} f(X(s-))(\sigma(s)\sigma^T(s))_{ij} ds + \int_t^T \int_{\|y\| \leq 1} y \cdot \nabla f(X(s-)) \widetilde{M}(dsdy) \\
&+ \int_t^T \int_{\mathbb{R}^n} (f(X(s-)+y) - f(X(s-)) - 1_{\|y\| \leq 1} y \cdot \nabla f(X(s-))) M(dsdy).
\end{aligned}
$$

Let E^x denote $E[\cdot|X(t-)=x]$, when taking expectations involving X or $E[\cdot|Y(t-)=x]$ when taking expectations involving Y. Taking expectations on both sides, we obtain

$$
\begin{aligned}
&E^x[f(X_T)] \\
=&f(x) + E^x \left[\int_t^T \nabla f(X(s-)) \cdot \beta(s) ds \right] + \frac{1}{2} E^x \left[\int_t^T \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} f(X(s-))(\sigma(s)\sigma^T(s))_{ij} ds \right] \\
&+ E^x \left[\int_t^T \int_{\mathbb{R}^n} (f(X(s-)+y) - f(X(s-)) - 1_{\|y\| \leq 1} y \cdot \nabla f(X(s-))) m(s,y) dsdy \right].
\end{aligned}
$$

We apply Fubini theorem and find that

$$
\begin{aligned}
&E^x[f(X_T)] \\
=&f(x) + \int_t^T E^x[\nabla f(X(s-)) \cdot \beta(s)] ds \\
&+ \frac{1}{2} \int_t^T \sum_{j,k=1}^n E^x \left[\frac{\partial^2}{\partial x_j \partial x_k} f(X(s-))(\sigma(s)\sigma^T(s))_{jk} \right] ds \\
&+ \int_t^T E^x [\int_{\mathbb{R}^n} (f(X(s-)+y) - f(X(s-)) - 1_{\|y\| \leq 1} y \cdot \nabla f(X(s-))) m(s,y)] dsdy.
\end{aligned}
$$

Using iterated expectations conditioned on $X(s-)$, we see that

$$\begin{aligned}
&E^x[f(X_T)] \\
=&f(x) + \int_t^T E^x[\nabla f(X(s-)) \cdot E^x[\beta(s)|X(s-)]]ds \\
&+ \frac{1}{2}\int_t^T \sum_{j,k=1}^n E^x\left[\frac{\partial^2}{\partial x_i \partial x_j} f(X(s-))E^x[(\sigma(s)\sigma^T(s))_{ij}|X(s-)]\right] ds \\
&+ \int_t^T E^x[E^x[\int_{\mathbb{R}^n} (f(X(s-)+y) - f(X(s-)) \\
&- 1_{\|y\|\leq 1} y \cdot \nabla f(X(s-)))m(s,y)dy|X(s-)]]ds.
\end{aligned}$$

Simplifying gives

$$\begin{aligned}
&E^x[f(X_T)] \\
=&f(x) + \int_t^T E^x[\nabla f(X(s-)) \cdot b(X(s-),s)]ds \\
&+ \frac{1}{2}\int_t^T \sum_{i,j=1}^n E^x[\frac{\partial^2}{\partial x_i \partial x_j} f(X(s-))a_{ij}(X(s-),s)]ds \\
&+ \int_t^T E^x[\int_{\mathbb{R}^n} (f(X(s-)+y) - f(X(s-)) - 1_{\|y\|\leq 1} y \cdot \nabla f(X(s-)))\nu(s,y,X(s-))dy]ds.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&E^x[f(X_T)] \\
=&f(x) + E^x\Big[\int_t^T \Big(\sum_{i=1}^n b_i(X(s-),s)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} a_{ij}(X(s-),s)\frac{\partial^2 f}{\partial x_i x_j} \\
&+ \int_{\mathbb{R}^n} \big(f(X(s-)+y) - f(X(s-)) - 1_{\|y\|\leq 1}\sum_{i=1}^n y_i \frac{\partial f}{\partial x_i})\nu(s,y,X(s-))dy\Big)ds\Big].
\end{aligned} \tag{4.3}$$

Similarly, suppose the process $Y(T)$ starts with $Y(t-) = x$, then

$$\begin{aligned}
&E^x[f(Y_T)] \\
=&f(x) + E^x\Big[\int_t^T \Big(\sum_{i=1}^n b_i(Y(s-),s)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} a_{ij}(Y(s-),s)\frac{\partial^2 f}{\partial x_i x_j} \\
&+ \int_{\mathbb{R}^n} \big(f(Y(s-)+y) - f(Y(s-)) - 1_{\|y\|\leq 1}\sum_{i=1}^n y_i \frac{\partial f}{\partial x_i})\nu(s,y,Y(s-))dy\Big)ds\Big].
\end{aligned}$$

Then, by Definition 4.2, we get

$$\hat{A}_s f(x) = \sum_{i=1}^{n} b_i(x,s)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x,s)\frac{\partial^2 f}{\partial x_i x_j} + \int_{\mathbb{R}^n} (f(x+y) - f(x) - 1_{\|y\|\leq 1}\sum_{i=1}^{n} y_i \frac{\partial f}{\partial x_i})\nu(s,y,x)dy, \tag{4.4}$$

and

$$E^x[f(Y(T))] = f(x) + E^x\Big[\int_t^T \hat{A}_s f(Y(s-))ds\Big]. \tag{4.5}$$

And also, by equation (4.3), we have

$$E^x[f(X(T))] = f(x) + E^x\Big[\int_t^T \hat{A}_s f(X(s-))ds\Big]. \tag{4.6}$$

Since $Y(t) = Y(t-)$ with probability one, we have

$$\lim_{T\downarrow t} \frac{E^x[f(Y(T))|Y(t-) = x]}{T-t} = (\hat{A}_t f)(x),$$

and similarly for $X(t)$,

$$\lim_{T\downarrow t} \frac{E^x[f(X(T))|X(t-) = x]}{T-t} = (\hat{A}_t f)(x),$$

since $X(t) = X(t-)$ with probability one. □

Now we can derive the forward equation.

Proposition 4.4. *(Forward equation)*

Let X be a semimartingale with decomposition (4.2) on $\mathbb{R}^n$ with generator

$$\hat{A}_t f(x) = \sum_{i=1}^{n} b_i(x,t)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x,t)\frac{\partial^2 f}{\partial x_i x_j} + \int_{\mathbb{R}^n} (f(x+y) - f(x) - 1_{\|y\|\leq 1}\sum_{i=1}^{n} y_i \frac{\partial f}{\partial x_i})\nu(t,y,x)dy,$$

for $f \in C_0^2(\mathbb{R}^n)$, and assume that the probability measure of $X(t)$ has a density $p_X(x,t)$, i.e.

$$E[f(X(t))] = \int_{\mathbb{R}^n} f(x)p_X(x,t)dx, \qquad f \in C_0^2.$$

Then $p_X(x,t)$ satisfies the forward equation

$$\hat{A_t}^* p = \frac{\partial p}{\partial t} \qquad \text{on } \mathbb{R}^n \times (0,\infty),$$

where $\hat{A_t}^*$ is the formal adjoint of $\hat{A}_t$ and is given by

$$\begin{aligned}\hat{A_t}^* g(x) =& \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2}{\partial x_i x_j}(a_{ij}(x,t)g) - \sum_{i=1}^{n} \frac{\partial}{\partial x_i}(b_i(x,t)g) \\ &+ \int_{\mathbb{R}^n} (g(x-y,t)\nu(t,y,x-y) - g(x)\nu(t,y,x) \\ &- 1_{\|y\|\le 1} \sum_{i=1}^{n} y_i \frac{\partial}{\partial x_i}(g\nu(t,y,x)))dy.\end{aligned}$$

Proof. For $f \in C_0^2(\mathbb{R}^n)$, we have

$$E[f(X(t))] = f(X(t_0)) + E\Big[\int_{t_0}^{t} \hat{A}_t f(X(s-))ds\Big].$$

Since $p_X(x,t)$ is the probability density function of $X(t)$, we can rewrite this equation as

$$\int_{\mathbb{R}^n} f(x)p_X(x,t)dx = f(X(t_0)) + \int_{t_0}^{t}\int_{\mathbb{R}^n} \hat{A}_t f(x)p(x,s)dxds.$$

Taking derivatives with respect to t of both sides of the preceding equation, we get

$$\begin{aligned}&\frac{\partial}{\partial t}\int_{\mathbb{R}^n} f(x)p_X(x,t)dx \\ =& \int_{\mathbb{R}^n} \big(\hat{A}_t f(x)\big)p_X(x,t)dx \\ =& \int_{\mathbb{R}^n} \sum_{i=1}^{n} b_i(x,t)\frac{\partial f}{\partial x_i}p_X(x,t) + \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x,t)\frac{\partial^2 f}{\partial x_i x_j}p_X(x,t) \\ &+ (f(x+y) - f(x) - 1_{\|y\|\le 1}\sum_{i=1}^{n} y_i \frac{\partial f}{\partial x_i})\nu(t,y,x)dyp_X(x,t)dx.\end{aligned}$$

Integrating by parts on the right hand side, we get

$$\begin{aligned}&\frac{\partial}{\partial t}\int_{\mathbb{R}^n} f(x)p_X(x,t)dx \\ =& \int_{\mathbb{R}^n} \Big(-\sum_{i=1}^{n} f(x)\frac{\partial}{\partial x_i}(b(x,t)p_X(x,t)) + \frac{1}{2}\sum_{i,j=1}^{n} f(x)\frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}(x,t)p_X(x,t))\Big)dx \\ &+ \int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\left(f(x+y) - f(x) - 1_{\|y\|\le 1}\sum_{i=1}^{n} y_i\frac{\partial f}{\partial x_i}\right)\nu(t,y,x)p_X(x,t)dydx.\end{aligned}$$

Consider the last integral in the preceding equation. For the first term, we need to shift x by $-y$, and for the last term, we integrate by parts once,

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\left(f(x+y)-f(x)-1_{\|y\|\le 1}\sum_{i=1}^n y_i\frac{\partial f}{\partial x_i}\right)\nu(t,y,x)p_X(x,t)dydx$$
$$=\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}(f(x)\nu(t,y,x-y)p(x-y,t)-f(x)\nu(t,y,x)p_X(x,t)$$
$$+f(x)1_{\|y\|\le 1}\sum_{i=1}^n y_i\frac{\partial}{\partial x_i}(\nu(t,y,x)p_X(x,t))dy)dx.$$

Thus,

$$\frac{\partial}{\partial t}\int_{\mathbb{R}^n} f(x)p_X(x,t)dx$$
$$=\int_{\mathbb{R}^n} f(x)\left(-\sum_{i=1}^n\frac{\partial}{\partial x_i}(b_i(x,t)p_X(x,t))+\frac{1}{2}\sum_{i,j=1}^n\frac{\partial^2}{\partial x_i\partial x_j}(a_{ij}(x,t)p_X(x,t))\right)dx$$
$$+\int_{\mathbb{R}^n}(\nu(t,y,x-y)p(x-y,t)-\nu(t,y,x)p_X(x,t)$$
$$+1_{\|y\|\le 1}\sum_{i=1}^n y_i\frac{\partial}{\partial x_i}(\nu(t,y,x)p_X(x,t))dy)dx$$
$$=\int_{\mathbb{R}^n} f(x)(\hat{A_t}^*p_X)(x,t)dx.$$

We obtain

$$\int_{\mathbb{R}^n} f(x)\frac{\partial}{\partial t}p_X(x,t)dx=\int_{\mathbb{R}^n}\left(\hat{A}_t f\right)(x)p_X(x,t)dx=\int_{\mathbb{R}^n} f(x)\left(\hat{A_t}^*p_X\right)(x,t)dx,$$

which we can write it in terms of the L^2 inner product,

$$\langle f,\frac{\partial p_X}{\partial t}\rangle=\langle\hat{A}_t f,p_X\rangle=\langle f,\hat{A_t}^*p_X\rangle,$$

where $\hat{A_t}^*$ is the formal adjoint of $\hat{A}_t$ and $p_X(x,t)$ is a solution of

$$\frac{\partial p_X}{\partial t}=\hat{A_t}^*p_X(x,t)$$

in the weak sense as in [12, Section 7.2]. □

Similarly, $p(y,t)$, the density function of the process $Y(t)$ also satisfies the equation

$$\frac{\partial p_Y}{\partial t}=\hat{A_t}^*p_Y(y,t)$$

4.3 Construction of fundamental solutions

In this section, we will consider the fundamental solutions of the forward equation defined by the generator $\hat{A}$.

The construction of fundamental solutions of parabolic differential equations is described by Friedman in [14, Sections 1.2 & 1.4] in the case of bounded domains $U \Subset R^n$ and extended to unbounded domains $U \subset R^n$ (including R^n) in [14, Section 1.6]. We follow his approach and construct fundamental solutions of the forward equation we derived in the previous section, noting that the forward equation is a partial integro-differential equation. Construction of fundamental solutions of integro-differential operators of this kind have also been described by Garroni and Menaldi [**?**, Theorem 4.3.6] for bounded domains in R^n and extended to the case of unbounded domains (in particular, R^n) by [**?**, §4.3.3].

Define L^* by the expression

$$L^* := \frac{\partial}{\partial t} - \hat{A_t}^*.$$

we simply this expression before we proceed:

$$\begin{aligned} L^*p(x,t) =& p_t(x,t) + \sum_{i=1}^{n}(b_i(x,t)p(x,t))_{x_i} - \frac{1}{2}\sum_{i,j=1}^{n}(a_{ij}(x,t)p(x,t))_{x_ix_j} \\ &- \int_{\mathbb{R}^n} p(x-y,t)\nu(t,y,x-y) - p(x,t)\nu(t,y,x) \\ &+ 1_{\|y\|\leq 1}\sum_{i=1}^{n} y_i(\nu(t,y,x)p(x,t))_{x_i}dy. \end{aligned}$$

This leads to

$$\begin{aligned} L^*p(x,t) =& p_t(x,t) - \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x,t)p_{x_ix_j}(x,t) \\ &+ \sum_{i=1}^{n}(b_i(x,t) - \frac{1}{2}\sum_{j=1}^{n} a_{ij}(x,t)_{x_j})p_{x_i}(x,t) + \sum_{i=1}^{n}(b_i(x,t)_{x_i} - \frac{1}{2}\sum_{j=1}^{n} a_{ij}(x,t)_{x_ix_j})p(x,t) \\ &- \int_{\mathbb{R}^n} p(x-y,t)\nu(t,y,x-y)dy + p(x,t)\int_{\mathbb{R}^n}\nu(t,y,x)dy \\ &- p(x,t)\int_{\mathbb{R}^n} 1_{\|y\|\leq 1}\sum_{i=1}^{n} y_i\nu(t,y,x)_{x_i}dy - p_{x_i}(x,t)\int_{\mathbb{R}^n} 1_{\|y\|\leq 1}\sum_{i=1}^{n} y_i\nu(t,y,x)dy. \end{aligned}$$

Denote

$$\begin{aligned}\hat{A}_{tij}(x,t) :=& \frac{1}{2}a_{ij}(x,t),\\ B_i(x,t) :=& b_i(x,t) - \frac{1}{2}\sum_{j=1}^{n} a_{ij}(x,t)_{x_j}\int_{\mathbb{R}^n} 1_{\|y\|\le 1}\sum_{i=1}^{n} y_i\nu(t,y,x)dy,\\ C(x,t) :=& \sum_{i=1}^{n}(b_i(x,t)_{x_i} - \frac{1}{2}\sum_{j=1}^{n} a_{ij}(x,t)_{x_ix_j})\\ &+\int_{\mathbb{R}^n}\nu(t,y,x) - 1_{\|y\|\le 1}\sum_{i=1}^{n} y_i\nu(t,y,x)_{x_i}dy.\end{aligned}$$

Then,

$$\begin{aligned}L^*p(x,t) =& p_t(x,t) - \sum_{i,j=1}^{n}\hat{A}_{tij}(x,t)p_{x_ix_j}(x,t) + \sum_{i=1}^{n} B_i(x,t)p_{x_i}(x,t)\\ &+ C(x,t)p(x,t) - \int_{\mathbb{R}^n} p(z,t)\nu(t,x-z,z)dz,\end{aligned}$$

Consider the equation

$$L^*p = 0, \tag{4.7}$$

where the coefficients A_{ij}, B_{ij}, C and n are defined on $U_T \equiv \bar{U}\times[0,T]$. The domain $U \subset \mathbb{R}^n$ can be unbounded and includes the important case $U = \mathbb{R}^n$. Throughout this section we assume:

(A3) L^* is parabolic in U.

(A4) The coefficients of L^* are continuous functions in U and, in addition, for all $(x,t),(x_0,t_0)\in U$, there exist constants $0 < M$, $0<\alpha<1$ such that

$$\begin{aligned}|A_{ij}(x,t) - A_{ij}(x-0,t_0)| &\le M\big(|x-x_0|^\alpha + |t-t_0|^{\alpha/2}\big),\\ |B_i(x,t) - B_i(x-0,t_0)| &\le M|x-x_0|^\alpha,\\ |C_(x,t) - C_(x-0,t_0)| &\le M|x-x_0|^\alpha.\end{aligned}$$

(A5) $\nu(t,y,x)$ is compactly supported, namely, there exists a constant M_1 such that $\nu(t,y,x)=0$ if $|x|\ge M_1$ or $|y|\ge M_1$.

We define the fundamental solutions of a forward equation follows, by analogy with the standard definition in [14, Section 1.1] of a fundamental solution for a parabolic partial differential equation.

Definition 4.3. *(Fundamental solutions of a forward equation)*
A fundamental solution of a forward equation $L^*p = 0$ in U_T is a function $p(x, t; y, \tau)$ defined for all $(x, t), (y, \tau) \in U_T$, $t > \tau$, which satisfies the following conditions:
(i) For fixed $(y, \tau) \in U_T$, it satisfies, as a function of (x, t), $x \in U$, $\tau < t < T$, the equation $L^*p = 0$;
(ii) For every continuous function f in $\bar{U}$ obeying the growth condition in Definition 3.1, if $x \in U$ then

$$\lim_{t \to \tau} \int_U p(x, t; y, \tau) f(y) dy = f(x).$$

For a Itô process, the operator L^* is

$$L^*p = p_t - \sum_{i,j=1}^{n} A_{ij}(x, t) p_{x_i x_j} + \sum_{i=1}^{n} B_i(x, t) p_{x_i} + C(x, t)p,$$

for a semimartingale, the operator L^* is

$$L^*p = p_t - \sum_{i,j=1}^{n} A_{ij}(x, t) p_{x_i x_j} + \sum_{i=1}^{n} B_i(x, t) p_{x_i} + C(x, t)p - \int_{\mathbb{R}^n} p(z, t) \nu(t, x - z, z) dz.$$

We adapt the parametrix method in [14, Section 1.2] for the construction of fundamental solutions of linear second-order parabolic PDEs to the construction of fundamental solutions of our linear second-order parabolic PIDE; see also [**?**, Chapter V] and [**?**, Section IV.11]. Existence of fundamental solutions for parabolic PIDEs of the kind examined in this thesis is also proved in [**?**], using related methods. The construction of fundamental solutions for our parabolic PIDE closely mirrors that of parabolic PDEs due to the assumption on the density defining the integral term.

First we introduce the function $G(x, t; y, \tau)$, for $t > \tau$,

$$G(x, t; y, \tau) = C(y, \tau) \frac{1}{(t - \tau)^{\frac{n}{2}}} \exp \left[\frac{-1}{4(t - \tau)} \sum_{i,j=1}^{n} A^{ij}(y, \tau)(x_i - y_i)(x_j - y_j) \right],$$

where

$$C(y, \tau) = \frac{1}{(2\sqrt{\pi})^n} [\det(A^{ij}(y, \tau))]^{\frac{1}{2}},$$

and $A^{ij}(x, t)$ is the inverse matrix to $A_{ij}(x, t)$. The function G is called the *parametrix*.

For each fixed (y,τ), the function $G(x,t;y,\tau)$ satisfies the following equation with "constant coefficients",

$$L_0^* p(x,t) := p_t - \sum_{i,j=1}^{n} A_{ij}(y,\tau) p_{x_i x_j} = 0,$$

and also satisfies the following proposition.

Proposition 4.5. *([14, Section 1.2 Theorem 1])*
Let f be a continuous function in U_T obeying the growth condition in Definition 3.1. Then

$$J(x,t,\tau) := \int_U G(x,t;y,\tau) f(y,\tau) dy$$

is continuous function in (x,t,τ), $x \in \bar{U}$, $0 \le \tau < T$ and

$$\lim_{t \to \tau} J(x,t,\tau) = f(x,t),$$

uniformly with respect to (x,t), $x \in S$, $0 < t \le T$, where S is any closed subset of U.

From now on, we use p to denote the fundamental solutions of the forward equation, because p is traditional for transitional probability density function. It follows from Proposition 4.5 that property (ii) in Definition 4.3 of the fundamental solution is also satisfied for $G = p$.

In order to construct a fundamental solution for $L^* p = 0$, we view L_0^* as a first approximation to L^* and we view G as the principal part of the fundamental solution G. We then try to find p in the form

$$p(x,t;y,\tau) = G(x,t;y,\tau) + \int_\tau^t \int_U G(x,t;\eta,\sigma) \Phi(\eta,\sigma;y,\tau) d\eta d\sigma,$$

where Φ is to be determined by the condition that p satisfies the equation ,

$$\begin{aligned} 0 &= L^* p(x,t;y,\tau) \\ &= L^* G(x,t;y,\tau) + L^* \int_\tau^t \int_U G(x,t;\eta,\sigma) \Phi(\eta,\sigma;y,\tau) d\eta d\sigma. \end{aligned}$$

We first consider the term $L^* G(x,t;y,\tau)$. For fixed (y,τ), $L_0^* G(x,t;y,\tau) = 0$, and so

$$L^*G(x,t;y,\tau) = L^*G(x,t;y,\tau) - L_0^*G(x,t;y,\tau) \tag{4.8}$$

$$= G_t - \sum_{i,j=1}^{n} A_{ij}(x,t)G_{x_ix_j} + \sum_{i=1}^{n} B_i(x,t)G_{x_i} + C(x,t)G \tag{4.9}$$

$$- \int_{\mathbb{R}^n} G(t,z;y,\tau)\nu(t,x-z,z)dz - G_t + \sum_{i,j=1}^{n} A_{ij}(y,\tau)G_{x_ix_j} \tag{4.10}$$

$$= - \sum_{i,j=1}^{n} (A_{ij}(x,t) - A_{ij}(y,\tau))G_{x_ix_j} + \sum_{i=1}^{n} B_i(x,t)G_{x_i} + C(x,t)G \tag{4.11}$$

$$- \int_{U} G(t,z;y,\tau)\nu(t,x-z,z)dz. \tag{4.12}$$

We need the following lemma to calculate

$$L^* \int_\tau^t \int_U G(x,t;\eta,\sigma)\Phi(\eta,\sigma;y,\tau)d\eta d\sigma.$$

Lemma 4.1. *([14, Section 1.3 Lemma 1])*

Let f be a continuous function in U obeying the growth condition in Definition 3.1 and locally Hölder continuous in $x \in U$, uniformly with respect to t, and

$$V(x,t) = \int_0^t \int_U G(x,t;\eta,\sigma)f(\eta,\sigma)d\eta d\sigma.$$

Then

$$\frac{\partial}{\partial x_i}V(x,t) = \int_0^T \int_U \frac{\partial}{\partial x_i}G(x,t;\eta,\sigma)f(\eta,\sigma)d\eta d\sigma,$$
$$\frac{\partial^2}{\partial x_i \partial x_j}V(x,t) = \int_0^T \int_U \frac{\partial^2}{\partial x_i \partial x_j}G(x,t;\eta,\sigma)f(\eta,\sigma)d\eta d\sigma,$$
$$\frac{\partial}{\partial t}V(x,t) = f(x,t) + \int_0^T \int_U \sum_{i,j=1}^{n} A_{ij}(\eta,\sigma)\frac{\partial^2}{\partial x_i \partial x_j}G(x,t;\eta,\sigma)f(\eta,\sigma)d\eta d\sigma.$$

If Φ is such that Lemma 4.1 applies to $f(x,t) := \Phi(x,t;y,\tau)$, then

$$V(x,t) = \int_\tau^t \int_U G(x,t;\eta,\sigma)\Phi(\eta,\sigma;y,\tau)d\eta d\sigma.$$

Consequently,

$$\begin{aligned}
L^*V(x,t) &= V_t - \sum_{i,j=1}^{n} A_{ij}(x,t)V_{x_ix_j} + \sum_{i=1}^{n} B_i(x,t)V_{x_i} + C(x,t)V \\
&\quad - \int_{\mathbb{R}^n} V(t,z)\nu(t,x-z,z)dz \\
&= \Phi(x,t;y,\tau) + \int_0^T \int_U \sum_{i,j=1}^{n} A_{ij}(\eta,\sigma)\frac{\partial^2}{\partial x_i \partial x_j} G(x,t;y,\tau)\Phi(\eta,\sigma;y,\tau)d\eta d\sigma \\
&\quad + \int_\tau^T \int_U - \sum_{i,j=1}^{n} A_{ij}(x,t)G_{x_ix_j} + \sum_{i=1}^{n} B_i(x,t)G_{x_i} + C(x,t)G d\eta d\sigma \\
&\quad - \int_\tau^T \int_U \int_U G(x,t;\eta,\sigma)\Phi(\eta,\sigma;y,\tau)\nu(t,x-z,z)d\eta d\sigma dz.
\end{aligned}$$

Therefore,

$$L^*V(x,t) = \Phi(x,t;y,\tau) + \int_\tau^T \int_U \Big\{ - [\sum_{i,j=1}^{n} A_{ij}(x,t) - A_{ij}(\eta,\sigma)]G_{x_ix_j} \tag{4.13}$$

$$+ \sum_{i=1}^{n} B_i(x,t)G_{x_i} + C(x,t)G d\eta d\sigma \tag{4.14}$$

$$- \int_U G(x,t;\eta,\sigma)\Phi(\eta,\sigma;y,\tau)\nu(t,x-z,z)dz\Big\} d\eta d\sigma \tag{4.15}$$

$$= \Phi(x,t;y,\tau) + \int_\tau^T \int_U L^*G(x,t;\eta,\sigma)\Phi(\eta,\sigma;y,\tau)d\eta d\sigma. \tag{4.16}$$

Combining (4.8) and (4.13) together, we get

$$\begin{aligned}
0 &= L^*p(x,t;y,\tau) \\
&= L^*G(x,t;y,\tau) + L^*V(x,t) \\
&= L^*G(x,t;y,\tau) + \Phi(x,t;y,\tau) + \int_\tau^T \int_U L^*G(x,t;\eta,\sigma)\Phi(\eta,\sigma;y,\tau)d\eta d\sigma.
\end{aligned}$$

Therefore,

$$-\Phi(x,t;y,\tau) = L^*G(x,t;y,\tau) + \int_\tau^T \int_U L^*G(x,t;\eta,\sigma)\Phi(\eta,\sigma;y,\tau)d\eta d\sigma. \tag{4.17}$$

Thus, for each fixed (y,τ), the function $\Phi(x,t;y,\tau)$ is a solution of a *Volterra integral equation* with singular kernel $L^*G(x,t;y,\tau)$.

Before we proceed to prove the existence of the fundamental solution, we give two useful lemmas. Lemma 4.2 is modeled after [14, Inequality (4.3), Section 1.4], but we need to evaluate an extra term, namely the integral term

$$\int_U G(t,z;y,\tau)\nu(t,x-z,z)dz$$

in the PIDE.

Lemma 4.2. *Let the notation be as above. Then,*

$$|L^*G(x,t;y,\tau)| \leq \frac{const}{(t-\tau)^\mu}\frac{1}{|x-y|^{n+2-2\mu-\alpha}} \tag{4.18}$$

where μ, α are constants, $1-\frac{\alpha}{2}<\mu<1$.

Proof. From the definition of L^*, we have

$$\begin{aligned}L^*G(x,t;y,\tau) = &-\sum_{i,j=1}^{n}(A_{ij}(x,t)-A_{ij}(y,\tau))G_{x_ix_j}+\sum_{i=1}^{n}B_i(x,t)G_{x_i}+C(x,t)G\\ &-\int_U G(t,z;y,\tau)\nu(t,x-z,z)dz,\end{aligned}$$

where

$$G(x,t;y,\tau)=C(y,\tau)\frac{1}{(t-\tau)^{\frac{n}{2}}}\exp\left[\frac{-1}{4(t-\tau)}\sum_{i,j=1}^{n}A^{ij}(y,\tau)(x_i-y_i)(x_j-y_j)\right],$$

and

$$C(y,\tau)=\frac{1}{(2\sqrt{\pi})^n}[\det(A^{ij}(y,\tau))]^{\frac{1}{2}}.$$

Here (A^{ij}) is the inverse matrix of (A_{ij}) and satisfies the ellipticity condition, that is, there exists a constant $\lambda>0$ such that

$$\sum_{i,j=1}^{n}A^{ij}(x_i-y_i)(x_j-y_j)\geq\lambda|x-y|^2.$$

For fixed (y,τ), if $0<\nu<\frac{n}{2}$, then

$$\begin{aligned}G(x,t;y,\tau) &\leq \frac{\text{const}}{(t-\tau)^{\frac{n}{2}}}\exp\left[-\frac{\lambda|x-y|^2}{4(t-\tau)}\right]\\ &=\frac{\text{const}}{(t-\tau)^{\nu}}\frac{1}{|x-y|^{n-2\nu}}\frac{|x-y|^{n-2\nu}}{(t-\tau)^{\frac{n}{2}-\nu}}\exp\left[-\frac{\lambda|x-y|^2}{4(t-\tau)}\right]\\ &\leq\frac{\text{const}}{(t-\tau)^{\nu}}\frac{1}{|x-y|^{n-2\nu}}.\end{aligned} \tag{4.19}$$

The last inequality is true because

$$\frac{|x-y|^{n-2\nu}}{(t-\tau)^{\frac{n}{2}-\nu}}\exp\left[-\frac{\lambda|x-y|^2}{4(t-\tau)}\right]=\left(\frac{|x-y|}{\sqrt{t-\tau}}\right)^{n-2\nu}\exp\left[-\frac{\lambda}{4}(\frac{|x-y|}{\sqrt{t-\tau}})^2\right]$$

is bounded as a function of $(x,t;y,\tau)$ in $U_T\times U_T\backslash D_T$, where $D_T=\{(x,t;t,\tau)\in U_T\times U_T : x=y, t=\tau\}$ as long as $0<\nu<\frac{n}{2}$ for any $n\geq 1$.

Let $\mu := \nu + (1 - \frac{\alpha}{2})$, then $1 - \frac{\alpha}{2} < \mu < 1$,

$$G(x,t;y,\tau) \leq \frac{\text{const}}{(t-\tau)^{\mu-(1-\frac{\alpha}{2})}} \frac{1}{|x-y|^{n+2-2\mu-\alpha}},$$
$$\leq \frac{\text{const}}{(t-\tau)^{\mu}} \frac{1}{|x-y|^{n+2-2\mu-\alpha}},$$

since $\frac{1}{(t-\tau)^{-(1-\frac{\alpha}{2})}}$ is bounded as $1 - \alpha/2 > 0$.

We assume that $C(x,t)$ is bounded as a function of $(x,t) \in U_T$, thus

$$C(x,t)G(x,t;y,\tau) \leq \frac{\text{const}}{(t-\tau)^{\mu}} \frac{1}{|x-y|^{n+2-2\mu-\alpha}}. \tag{4.20}$$

Similarly, we can apply this trick to G_{x_i} and $G_{x_i x_j}$ and get

$$\frac{\partial}{\partial x_i} G(x,t;y,\tau) \leq \frac{\text{const}|x-y|}{(t-\tau)^{\frac{n}{2}+1}} \exp\left[-\frac{\lambda|x-y|^2}{4(t-\tau)}\right],$$
$$\leq \frac{\text{const}}{(t-\tau)^{\mu}} \frac{1}{|x-y|^{n+2-2\mu-\alpha}},$$

$$\frac{\partial^2}{\partial x_i \partial x_j} G(x,t;y,\tau) \leq \frac{\text{const}|x-y|^2}{(t-\tau)^{\frac{n}{2}+2}} \exp\left[-\frac{\lambda|x-y|^2}{4(t-\tau)}\right],$$
$$\leq \frac{\text{const}}{(t-\tau)^{\mu}} \frac{1}{|x-y|^{n+2-2\mu}}.$$

By assumption $(A4)$, $B(x,t)$ is bounded and $|A_{ij}(x,t) - A_{ij}(y,\tau)| \leq \text{const}|x-y|^{\alpha}$, we get

$$B(x,t)\frac{\partial}{\partial x_i} G(x,t;y,\tau) \leq \frac{\text{const}}{(t-\tau)^{\mu}} \frac{1}{|x-y|^{n+2-2\mu-\alpha}}, \tag{4.21}$$

$$(A_{ij}(x,t) - A_{ij}(y,\tau))\frac{\partial^2}{\partial x_i \partial x_j} G(x,t;y,\tau) \leq \frac{\text{const}}{(t-\tau)^{\mu}} \frac{1}{|x-y|^{n+2-2\mu-\alpha}}. \tag{4.22}$$

Now we consider the integral

$$\int_U G(z,t;y,\tau)\nu(t,x-z,z)dz.$$

Because ν is compactly supported as a function of $(x,z) \in \mathbb{R}^n \times \mathbb{R}^n$, for $|x-y| \leq 2M_1$ we have

$$\int_U G(z,t;y,\tau)\nu(t,x-z,z)dz \leq \text{const} \int_U G(z,t;y,\tau)dz$$
$$\leq \text{const} \int_U \frac{1}{(t-\tau)^{\frac{n}{2}}} \exp[-\frac{\lambda|z-y|^2}{4(t-\tau)}]dz$$
$$\leq \text{const},$$

where the last estimate follows because the function

$$\frac{1}{(t-\tau)^{\frac{n}{2}}}\exp\left[-\frac{\lambda|x-y|^2}{4(t-\tau)}\right]$$

is integrable.

When $|x-y|>2M_1$, by changing the variables $x-z$ and z, we get

$$\int_U G(z,t;y,\tau)\nu(t,x-z,z)dz=-\int_U G(x-z,t;y,\tau)\nu(t,z,x-z)dz.$$

Since ν is bounded function with compact support, we have

$$\int_U G(z,t;y,\tau)\nu(t,x-z,z)dz\le \text{const}\int_{-M_1}^{M_1} G(x-z,t;y,\tau)dz.$$

By equation (4.19), we get

$$\int_{-M_1}^{M_1} G(x-z,t;y,\tau)dz\le \text{const}\int_{-M_1}^{M_1}\frac{1}{(t-\tau)^{\frac{n}{2}}}\exp[-\frac{\lambda|x-y-z|^2}{4(t-\tau)}]dz$$

Because $|x-y|>2M_1$ and $|z|<M_1$, so

$$|x-y-z|\ge|x-y|-|z|\ge|x-y|-M_1\ge|x-y|-\frac{1}{2}|x-y|=\frac{1}{2}|x-y|,$$

$$-|x-y-z|^2\le\frac{1}{4}|x-y|^2,$$

thus

$$\int_{-M_1}^{M_1}\frac{1}{(t-\tau)^{\frac{n}{2}}}\exp[-\frac{\lambda|x-y-z|^2}{4(t-\tau)}]dz\le\int_{-M_1}^{M_1}\frac{1}{(t-\tau)^{\frac{n}{2}}}\exp[-\frac{\frac{\lambda}{4}|x-y|^2}{4(t-\tau)}]dz.$$

Using the same trick applied in equation (4.19), we get

$$\frac{1}{(t-\tau)^{\frac{n}{2}}}\exp[-\frac{\frac{\lambda}{2}|x-y|^2}{4(t-\tau)}]\le\text{const}\frac{1}{(t-\tau)^{\frac{n}{2}}}\exp[-\frac{\frac{\lambda}{2}|x-y|^2}{4(t-\tau)}].$$

Therefore,

$$\int_U G(z,t;y,\tau)\nu(t,x-z,z)dz$$

also satisfies the inequality

$$\int_U G(z,t;y,\tau)\nu(t,x-z,z)dz\le\frac{\text{const}}{(t-\tau)^{\mu}}\frac{1}{|x-y|^{n+2-2\mu-\alpha}},\tag{4.23}$$

Adding inequalities (4.20), (4.21), (4.22), (4.23) together, we obtain

$$|L^*G(x,t;y,\tau)|\le\frac{\text{const}}{(t-\tau)^{\mu}}\frac{1}{|x-y|^{n+2-2\mu-\alpha}},$$

when $1-\frac{\alpha}{2}<\mu<1$.

□

The following Lemma is based on [14, Lemma 2, Section 1.4] which only holds for bounded domain. We extend it to the case of an unbounded domain.

Lemma 4.3. *Suppose that U is a domain in $\mathbb{R}^n$, which could be $\mathbb{R}^n$, $0 < \alpha < n$, $0 < \beta < n$, and $\alpha + \beta > n$ then for any $x \in U, z \in U, x \neq z$, we have*

$$\int_U \frac{1}{|x-y|^\alpha |y-z|^\beta} dy \leq const \cdot |x-z|^{n-\alpha-\beta}.$$

Proof. Let $U_1 = \{y \in U | s.t. |y-z| < \frac{|x-z|}{2}\}$, $U_2 = \{y \in U | s.t. |y-x| < \frac{|x-z|}{2}\}$ and $U_3 = U \backslash U_1 \cup U_2$.

First, we estimate the integral on U_1:

$$\begin{aligned} \int_{U_1} \frac{1}{|x-y|^\alpha |y-z|^\beta} dy &\leq \int_{U_1} \frac{2^\alpha}{|x-z|^\alpha |y-z|^\beta} dy \\ &\leq \omega_{n-1} d \frac{2^\alpha}{|x-z|^\alpha} \int_0^{\frac{|x-z|}{2}} r^{n-1-\beta} dr \\ &= const \cdot |x-z|^{n-\alpha-\beta}. \end{aligned}$$

Similarly,

$$\int_{U_2} \frac{1}{|x-y|^\alpha |y-z|^\beta} dy \leq const \cdot |x-z|^{n-\alpha-\beta}.$$

Now we estimate $I_3 := \int_{U_3} \frac{1}{|x-y|^\alpha |y-z|^\beta} dy$. First we notice

$$\frac{1}{|x-y|^\alpha |y-z|^\beta} \leq \frac{1}{|x-y|^{\alpha+\beta}} + \frac{1}{|z-y|^{\alpha+\beta}}.$$

Thus

$$\begin{aligned} I_3 &\leq \int_{U_3} \left[\frac{1}{|x-y|^{\alpha+\beta}} + \frac{1}{|z-y|^{\alpha+\beta}} \right] dy \\ &\leq 2\omega_{n-1} \int_{\frac{2}{|x-z|}}^{\infty} r^{n-1-\alpha-\beta} dr \\ &= const \cdot |x-z|^{n-\alpha-\beta}, \quad \text{for} \quad \alpha+\beta > n. \end{aligned}$$

□

The following result closely follows the construction of Φ in [14, Section 1.4] — see the proofs of [14, Section 1.4, Theorems 7 & 8]; this is extended to unbounded domains in [14, Section 1.6] — see the proofs of [14, Section 1.6, Theorems 10 & 11].

Theorem 4.2. *There exists a solution* Φ *of equation (4.17) of the form*

$$-\Phi(x,t;y,\tau) = \sum_{k=1}^{\infty} (L^*G)_k(x,t;y,\tau), \tag{4.24}$$

where

$$(L^*G)_1 := L^*G,$$
$$(L^*G)_{k+1}(x,t;y,\tau) := \int_\tau^T \int_U [L^*G(x,t;\eta,\sigma)](L^*G)_k(\eta,\sigma;y,\tau)d\eta d\sigma.$$

Proof. Using Lemma 4.2 and Lemma 4.3, we get [1]

$$\begin{aligned}
|(L^*G)_2(x,t;y,\tau)| &= \int_\tau^T \int_U [L^*G(x,t;\eta,\sigma)](L^*G)(\eta,\sigma;y,\tau)d\eta d\sigma \\
&\le \text{const} \int_\tau^T \int_U \frac{1}{(t-\sigma)^\mu} \frac{1}{|x-\eta|^{n+2-2\mu-\alpha}} \frac{1}{(\sigma-\tau)^\mu} \frac{1}{|\eta-y|^{n+2-2\mu-\alpha}} d\eta d\sigma \\
&\le \text{const} \int_\tau^T \frac{1}{(t-\sigma)^\mu} \frac{1}{(\sigma-\tau)^\mu} d\sigma \int_U \frac{1}{|x-\eta|^{n+2-2\mu-\alpha}} \frac{1}{|\eta-y|^{n+2-2\mu-\alpha}} d\eta \\
&\le \text{const} \frac{1}{(t-\tau)^{2\mu-1}} \frac{1}{|x-y|^{n+2(2-2\mu-\alpha)}},
\end{aligned}$$

when $2\mu > 1$ and $n + 2(2-2\mu-\alpha) > 0$. Since $1 - \frac{\alpha}{2} < \mu < 1$, the singularity of (LG_2) is weaker that that of LG. Similarly, we can proceed to evaluate $(L^*G)_3$,

$$\begin{aligned}
&|(L^*G)_3(x,t;y,\tau)| \\
&= \int_\tau^T \int_U [L^*G(x,t;\eta,\sigma)]_2(L^*G)(\eta,\sigma;y,\tau)d\eta d\sigma \\
&\le \text{const} \int_\tau^T \frac{1}{(t-\sigma)^{2\mu-1}} \frac{1}{(\sigma-\tau)^\mu} d\sigma \int_U \frac{1}{|x-\eta|^{n+2(2-2\mu-\alpha)}} \frac{1}{|\eta-y|^{n+2-2\mu-\alpha}} d\eta \\
&\le \text{const} \frac{1}{(t-\tau)^{3\mu-2}} \frac{1}{|x-y|^{n+3(2-2\mu-\alpha)}},
\end{aligned}$$

when $3\mu > 2$ and $n + 3(2-2\mu-\alpha) > 0$.

We know after finite steps, we arrive at some k_0 for which $k_0\mu < k_0 - 1$, and $n + k_0(2-2\mu-\alpha) < 0$, thus

$$|(L^*G)_{k_0}(x,t;y,\tau)| \le \text{const}.$$

From k_0, we proceed to prove by induction on m, assume that

$$|(L^*G)_{m+k_0}(x,t;y,\tau)| \le C_0 \frac{[C(t-\tau)^{1-\mu}]^m}{\Gamma(1+(1-\mu)m)},$$

---

[1] See Remark 4.1 for additional details for the case of unbounded domains $U \subset R^n$.

where C_0, C are constants and $\Gamma(t)$ is the gamma function. For $m = 0$, this follows from $|(L^*G)_{k_0}(x,t;y,\tau)| \leq \text{const}$. Assuming now that it holds for some integer $m \geq 0$, and using Lemma 4.2 we get

$$|(L^*G)_{m+1+k_0}(x,t;y,\tau)| \leq \text{const}.C_0 \frac{C^m}{\Gamma((1-\mu)m+1)} \int_\tau^t (t-\sigma)^{-\mu}(\sigma-\tau)^{(1-\mu)m} d\sigma.$$

Substituting $\rho = \frac{\sigma-\tau}{t-\tau}$ into the preceding expression and using the formula

$$\int_0^1 (1-\rho)^{a-1}\rho^{b-1} d\rho = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},$$

we obtain

$$\begin{aligned}\int_\tau^t (t-\sigma)^{-\mu}(\sigma-\tau)^{(1-\mu)m} d\sigma &= \int_0^1 (t-\tau)^{1-\mu}(m+1)\rho^{1-\mu}m(1-\rho)^{-\mu} d\sigma \\ &= (t-\tau)^{1-\mu}(m+1)\frac{\Gamma(1-\mu)\Gamma(1+m(1-\mu))}{\Gamma(1+(1-\mu)(m+1))}.\end{aligned}$$

Thus

$$|(L^*G)_{m+1+k_0}(x,t;y,\tau)| \leq C_0 \frac{[C(t-\tau)^{1-\mu}]^{m+1}}{\Gamma(1+(1-\mu)(m+1))},$$

and the induction holds for $m+1$.

It follows that

$$\begin{aligned}-\Phi(x,t;y,\tau) &= \sum_{k=1}^{\infty}(L^*G)_k(x,t;y,\tau) \\ &= \sum_{k=1}^{k_0}(L^*G)_k(x,t;y,\tau) + \sum_{m=1}^{\infty}(L^*G)_{k_0+m}(x,t;y,\tau) \\ &\leq \frac{\text{const}}{(t-\tau)^\mu}\frac{1}{|x-y|^{n+2-2\mu-\alpha}},\end{aligned}$$

and the series converges. □

From Theorem 4.2, it follows that the series expansion (4.24) for $\Phi(x,t;y,\tau)$ converges and that integral term in (4.17) is equal to

$$\sum_{k=1}^{\infty}\int_\tau^T \int_U L^*G(x,t;\eta,\sigma)\cdot(L^*G)_k(\eta,\sigma;y,\tau) d\eta d\sigma.$$

Therefore,

$$p(x,t;y,\tau) = G(x,t;y,\tau) + \int_\tau^t \int_U G(x,t;\eta,\sigma)\Phi(\eta,\sigma;y,\tau) d\eta d\sigma,$$

satisfies (4.7), so property (i) in Definition 4.3 holds. By Proposition 4.5, property (ii) also holds. Therefore $p(x,t;y,\tau)$ is a fundamental solution of (4.7). Compare the statements and proofs of [14, Section 1.4, Theorem 8] for bounded domains and [14, Section 1.6, Theorem 10] for unbounded domains.

Remark 4.1. *Friedman notes in [14, Section 1.6] that the construction of the fundamental solution, $p(x,t;y,\tau)$, extends from the case of a bounded domain $U \Subset R^n$ to an unbounded domain $U \subset R^n$ and in particular $U = R^n$. We briefly summarize one approach to the changes for unbounded domains here and refer the reader to standard references for further details [**?**, Chapter V] and [**?**, Section IV.11]. The estimate in Lemma 4.2 is replaced [**?**, Chapter V, Equation (3.19)], [**?**, Chapter 4, Equation (11.17)] by the better behaved*

$$|L^*G(x,t;y,\tau)| \leq c(t-\tau)^{\frac{1}{2}(\alpha-n-2)} \exp\left(-C\frac{|x-y|^2}{t-\tau}\right),$$

where $\alpha \in (0,1)$ is the Hölder constant, as before, and C, c are positive constants; Lemma 4.3 will not be used. This estimate is standard when no integral term appears in the definition of L^*, while the proof of Lemma 4.2 shows that the integral term in $L^*G(x,t;y,\tau)$ also obeys this estimate; indeed, the proof of Lemma 4.2 yields

$$\begin{aligned}\int_U G(z,t;y,\tau)\nu(t,x-z,z)dz &\leq c(t-\tau)^{-\frac{n}{2}} \exp\left(-C\frac{|x-y|^2}{t-\tau}\right) \\ &\leq c(t-\tau)^{\frac{1}{2}(\alpha-n-2)} \exp\left(-C\frac{|x-y|^2}{t-\tau}\right),\end{aligned}$$

using $\frac{1}{2}(\alpha-2) \in (-1,-\frac{1}{2})$ and $0 \leq t-\tau \leq T$, and this bound replaces (4.23).

Next, the estimate appearing in the proof of Theorem 4.2 for the term $(L^*G)_k(x,t;y,\tau)$ in the infinite series defining $\Phi(x,t;y,\tau)$, obtained by the iterative method of solving the Volterra integral equation (4.17), *is replaced by [**?**, Chapter V, Equation (3.22)], [**?**, Chapter 4, Equation (11.25)]*

$$|(L^*G)_k(x,t;y,\tau)| \leq c^k \left(\frac{\pi}{C}\right)^{\frac{n(k-1)}{2}} \frac{\Gamma^k(\alpha/2)}{\Gamma(k\alpha/2)}(t-\tau)^{-\frac{1}{2}(k\alpha-n-2)} \exp\left(-C\frac{|x-y|^2}{t-\tau}\right),$$

for $k \geq 1$, where $(L^*G)_1 := L^*G$. These estimates for $(L^*G)_k(x,t;y,\tau)$ ensure uniform convergence of the series in the statement of Theorem 4.2 for $t-\tau > 0$ and yields the *estimate*

$$|\Phi(x,t;y,\tau)| \le c(t-\tau)^{\frac{1}{2}(\alpha-n-2)} \exp\left(-C\frac{|x-y|^2}{t-\tau}\right),$$

just as in [?, Chapter 4, Equation (11.26)].

4.4 Existence and uniqueness of weak solutions

In this section, we will show the existence and uniqueness of weak solutions of the partial integral equation (4.7)

$$L^* p = 0.$$

Let U be an open subset of $\mathbb{R}^n$, and set $U_T = U \times (0,T]$. U can be unbounded in $\mathbb{R}^n$, and the special case $U = \mathbb{R}^n$ is of particular importance.

Let

$$Au = -\sum_{i,j=1}^{n} (a_{ij}(x,t)u_{x_i})_{x_j} + \sum_{i=1}^{n} (b_i(x,t)u)_{x_i} + c(x,t)u + \int_U u(z,t)\nu(t,x-z,z)dz,$$

where $\nu(t,x,z) : [0,T] \times U \times U \mapsto \mathbb{R}$.

We will study the following parabolic equation with initial and boundary conditions

$$\begin{aligned} u_t + Au =& f \quad \text{in} \;\; U_t \qquad (4.25)\\ u =& 0 \quad \text{on} \;\; \partial U \times [0,T] \\ u =& g \quad \text{on} \;\; U \times \{t=0\} \end{aligned}$$

We assume that the coefficients of L satisfy the following conditions (A5):

$$\sum_{i,j=1}^{n} a_{ij}(x,t)\xi_i\xi_j \ge \theta|\xi|^2 \quad \text{for all} \;\; (x,t) \in U_T, \xi \in \mathbb{R}^n \qquad (4.26)$$

$$a_{ij}, b_i, c \in L^\infty(U_T) \qquad (4.27)$$

$$\nu \in L^\infty(0,T;L^2(U \times U)) \qquad (4.28)$$

$$f \in L^2(U_T) \qquad (4.29)$$

$$g \in L^2(U) \qquad (4.30)$$

Remark 4.2. *$n \in L^\infty(0,T;L^2(U \times U))$ means*

$$\|\nu(t,\cdot,\cdot)\|_{L^2(U\times U)} := \int_U \int_U \nu^2(t,x,z)dxdz$$

is bounded by a finite constant for a.e. $t \in [0, T]$.

To apply a theorem in [26, Section 3.2], we need the following lemma.

Lemma 4.4. *Suppose* a_{ij}, b_i, c *satisfy (A5). Then there exist positive constants* λ *and* C *depending only on the coefficients of* L *such that*

$$(Lu, u) + \lambda\|u\|_{L^2(U)} \geq C\|u\|_{H_0^1(U)}$$

for a.e. $t \in [0, T]$ *and all* $u \in H^1(U)$.

Proof. Since a_{ij} satisfies the elipticity condition (4.27), we have

$$\Big|\int_U a_{ij}(x,t)u_{x_i}u_{x_j}dx\Big| \geq \theta\|\nabla u\|^2_{L^2(U)}, \tag{4.31}$$

for a.e. $t \in [0, T]$ and some positive constant θ.

Since $b_i \in L^\infty(U_T)$, then for a.e $t \in [0, T]$ the Cauchy inequality yields

$$\Big|\int_U b_i(t,x)u(t,x)u_{x_i}(x)\Big| \leq M\|u\|_{L^2(U)}\|u\|_{H^1(U)}.$$

Then for any $\varepsilon \geq 0$, there exists $M_\varepsilon > 0$ such that

$$\int_U b_i(t,x)u(t,x)u_{x_i}(x) \geq -M\|u\|_{L^2(U)}\|u\|_{L^2(H^1(U))} \tag{4.32}$$

$$\geq -M_\varepsilon\|u\|^2_{L^2(U)} - \varepsilon\|u\|^2_{H^1(U)}. \tag{4.33}$$

Because $c \in L^\infty(U_T)$, then for a.e $t \in [0, T]$, there exists a constant M_1 such that

$$\Big|\int_U cu^2(t,x)dx\Big| \leq M_1\|u\|^2_{L^2(U)}. \tag{4.34}$$

Since

$$\int_{U\times U} n^2(t,x,z)dxdz$$

is bounded by a finite number for a.e. $t \in [0,T]$, the Cauchy inequality gives

$$\begin{aligned}
&\Big|\int_U\int_U u(z)u(x)\nu(t,x-z,z)dxdz\Big| \qquad (4.35)\\
\leq&(\int_U u^2(x)dx)^{\frac{1}{2}}(\int_U(\int_U u(z)\nu(t,x-z,z)dz)^2dx)^{\frac{1}{2}}\\
\leq&\|u\|_{L^2(U)}(\int_u(\int_U u^2(z)dz\int_U \nu^2(t,x-z,z)dz)dx)^{\frac{1}{2}}\\
=&\|u|^2_{L^2(U)}\int_U\int_U \nu^2(t,x-z,z)dxdz\\
\leq&M_2\|u\|^2_{L^2(U)}.
\end{aligned}$$

for a.e $t \in [0,T]$ and some constant M_2.

By combining (4.31), (4.32), (4.34) and (4.35), we obtain

$$\begin{aligned}
< Lu,u >_{L^2(U)} =& \int_U \{a_{ij}(x,t)u_{x_i}u_{x_j} + b_i(t,x)u(t,x)u_{x_i}(x) + cu^2\}dx\\
&+\int_U\int_U u(z)u(x)\nu(t,x-z,z)dzdx\\
&\geq \theta\|\nabla u\|^2_{L^2(U)} - M_\varepsilon\|u\|^2_{L^2(U)} - \varepsilon\|u\|^2_{H^1(U)} - M_1\|u\|_{L^2(U)} - M_2\|u\|^2_{L^2(U)}.
\end{aligned}$$

Setting $\varepsilon := \theta/2$, $\lambda := M_\varepsilon + M_1 + M_2$, we obtain

$$< Lu,u >_{L^2(U)} +\lambda\|u\|_{L^2(U)} \geq C\|u\|_{H^1_0(U)},$$

where $C := \theta/2$. This completes the proof. □

Let V be a separable Hilbert space with dual V'; then $L^2(0,T;V)$ is a Hilbert space with dual $L^2(0,T;V')$. Assume that for each $t \in [0,T]$ we are given a continuous bilinear form $a(t;\cdot,\cdot)$ on V or, equivalently, an operator $\mathcal{A}(t) \in \mathcal{L}(V,V')$,

$$\mathcal{A}(t)u(v) = a(t;u,v), \quad u,v \in V, \quad t \in [0,T],$$

such that for each pair $u,v \in V$ the function $a(\cdot;u,v)$ is in $L^\infty(0,T;\mathbb{R})$. Assume H is a Hilbert space identified with its dual and that the embedding $V \hookrightarrow H$ is dense and continuous;hence $H \subset V'$ by restriction. Finally, let $f \in L^2(0,T;V')$ and $u_0 \in H$ be given.

Consider the abstract Cauchy problem

$$u \in L^2(0,T;V): \ u' + Au = f \text{ in } L^2(0,T;V'), \quad u(0) = u_0$$

where the separable Hilbert spaces $V \hookrightarrow H \hookrightarrow V'$, bounded and measurable operators $A(t) : V \mapsto V'$, and $f \in L^2(0,T;V')$, $u_0 \in H$ are given as above.

Proposition 4.6. *[26, Proposition 2.3]*

Assume the operators are uniformly coercive: there is a $c > 0$ such that

$$A(t)v(v) \geq c\|v\|_V^2, v \in V, t \in [0,T].$$

Then there exists a unique solution of the Cauchy problem, and it satisfies

$$\|u\|_{L^2(0,T;V)}^2 \leq (1/c)^2(\|f\|_{L^2(0,T;V')}^2 + |u_0|_H^2).$$

This result can be extended. $u \in H$ if and only if $v \in H$ where $v(t) = e^{-\lambda t}u(t)$, $0 \leq t \leq T$ and u is a solution of the proceeding Cauchy problem exactly when v is the corresponding solution of the problem

$$v \in H : u' + (A(\cdot) + \lambda I)v = e^{-\lambda t}f(t), v(0) = u_0.$$

Corollary 4.1. *A sufficient condition for existence by Proposition (4.6) is that there exist a $\lambda \in \mathbb{R}$ and $c > 0$ such that*

$$A(t)v(v) + \lambda|v|_H^2 \geq c\|v\|_V^2, v \in V, t \in [0,T].$$

Similarly, uniqueness is obtained from such an estimate, even with $c = 0$.

Theorem 4.3. *If the coefficients of A satisfy (A4) and (A5), then there exists a unique weak solution of (4.25).*

Proof. We choose $H = L^2(U)$, $V = H_0^1(U)$ for $U \subsetneq \mathbb{R}^n$ or $V = H^1(\mathbb{R}^n)$ for $U = \mathbb{R}^n$, then $V' = H^{-1}(U)$. Lemma 4.4 implies

$$\|u\|_{L^2(U)}^2 + \lambda\|u\|_{L^2(U)} \geq C\|u\|_{H_0^1(U)}.$$

We obtain the desired result by Proposition 4.6. $\square$

4.5 Uniqueness of fundamental solutions

Theorem 4.4. *There exists a unique solution to (4.7).*

Proof. Assume that there exist two fundamental solutions $p(x,t)$ and $\tilde{p}(x,t)$. Given any initial condition $g \in C_0^\infty(U)$, suppose the following parabolic equation with initial and boundary conditions

$$\begin{aligned} &u_t + Lu = f \quad \text{in } U_T, \\ &u = 0 \quad \text{on } \partial U \times [0,T], \\ &u = g \quad \text{on } U \times \{t = 0\}, \end{aligned}$$

has two solutions which can be expressed in terms of fundamental solutions $p(x,t)$ and $\tilde{p}(x,t)$

$$\begin{aligned} u(x,t) &= \int_U p(x-z,t)g(z)dz, \\ \tilde{u}(x,t) &= \int_U \tilde{p}(x-z,t)g(z)dz. \end{aligned}$$

According to Theorem 4.3, the functions u and $\tilde{u}$ are equal for a.e. $t \in [0,T]$. Thus

$$\int_U (p(x-z,t) - \tilde{p}(x-z,t))g(z)dz = 0 \quad \text{for a.e. } (x,t) \in U_T,$$

for every $g \in C_0^\infty(U)$. This implies

$$p(x,t) = \tilde{p}(x,t) \quad \text{for a.e. } (x,t) \in U_T.$$

□

The partial integro-equation,

$$L^*p = p_t - \sum_{i,j=1}^n A_{ij}(x,t)p_{x_ix_j} + \sum_{i=1}^n B_i(x,t)p_{x_i} + C(x,t)p - \int_{\mathbb{R}^n} p(z,t)\nu(t,x-z,z)dz = 0 \tag{4.36}$$

has a unique fundamental solution. Since the marginal density function $p_X(x,t)$ of the semimartingale $X(t)$ and the the marginal density function $p_Y(y,t)$ of the mimicking process $Y(t)$ both satisfy (4.36), the uniqueness of the fundamental solution of (4.36) implies that $X(t)$ and $Y(t)$ have the same marginal distributions.

This is an extension of Theorem 3.3 for the case of a semimartingale. It is also justified in [25, Theorem 3.4.2].

Chapter 5

Markov Processes and Pseudo-Differential Operators

When we have a partial integro-differential equation, it is natural to investigate it using pseudo differential operators. However, in the forward equation we derive, the integral term is $\int_{\mathbb{R}^n} p(z,t)\nu(t,x-z,z)dz$ which is not a standard convolution. So we cannot apply the general theory of pseudo-differential operators.

In this chapter, we review operator semigroups, Feller processes and discuss how research of pseudo differential operators arises in the martingale problem. To end this chapter, we indicate an area of further study. We plan to investigate the generator A of a semimartingale $X(t)$. We first want to show that A is a pseudo-differential operator with a symbol $a(t,x,\xi)$. In principle we could check that $a(t,x,\xi)$ satisfies certain conditions in [6, Theorem 4.2] which should imply the existence of a Markov process with generator A. We believe that this could lead to a new proof of the mimicking theorems, but this appears to be a challenging problem and we leave it for future research.

5.1 Operator semigroups and Feller processes

In this section, we give a brief introduction to operator semigroups and their generators from a probabilistic perspective. We outline the relationship between operator semigroups and Feller processes.

Definition 5.1. *(Operator semigroups [18] Definition 4.1.1)*

Let $(X, \|\cdot\|_X)$ be a Banach space. Then a one parameter family of bounded linear *operators* $(T_t)_{t\geq 0} \in \mathcal{L}(X, X)$ *is called an operator semigroup if*

$$T_{t+s} = T_t \circ T_s, \quad \text{for all} \;\; s, t \geq 0, \qquad (5.1)$$

$$T_0 = I.$$

We call $(T_t)_{t\geq 0}$ *strongly continuous if*

$$\lim_{t\mapsto 0} \|T_t u - u\|_X = 0,$$

for all $u \in X$*.*

The semigroup $(T_t)_{t\geq 0}$ *is called a contraction semigroup, if for all* $t \geq 0$*,*

$$\|T_t\| \leq 1$$

holds, that is, if each of the operators T_t *is a contraction. As usual,* $\|T_t\|$ *denotes the operator norm.*

It is easy to see that (5.1) corresponds to the exponential Cauchy functional equation

$$g(s+t) = g(s)g(t), \qquad g(0) = 1,$$

where $g(\cdot)$ is a nonnegative function from $\mathbb{R}$ to $\mathbb{R}$. The solution to the exponential Cauchy functional equation is the family of exponential functions $g(t) = e^{\alpha t}$, $\alpha \in \mathbb{R}$. However, this family represents all possible solutions only if an additional assumption of continuity is made. In fact, the assumption that $g(t)$ is continuous from the right in the origin is already sufficient to make the functions $g(t) = e^{\alpha t}$ the only solutions. We now introduce a similar assumption for the operator semigroup defined by (5.1). Now we want to show that by analogy to the Cauchy equation, an operator semigroup can be represented in the form $T_t = e^{tA}$ for a suitable operator A.

Definition 5.2. *(Generators of semigroups [18] Definition 4.1.11)*
Let $(T_t)_{t\geq 0}$ *be a strongly continuous semigroup of operators on a Banach space* $(X, \|\cdot\|_X)$*. The generator* A *of* $(T_t)_{t\geq 0}$ *is defined by*

$$(Au)(x) = \lim_{t\to 0^+} \frac{T_t u(x) - u(x)}{t}, \qquad (5.2)$$

with domain

$$D(A) = \Big\{ u \in X \Big| \lim_{t\to 0^+} \frac{T_t u(x) - u(x)}{t} \text{ exists as strong limits.} \Big\}.$$

While the infinitesimal generator A is defined as the right-hand derivative of T_t at 0, the derivative of T_t at any point can be calculated by

$$\frac{d}{dt}(T_t u)(x) = \lim_{h \mapsto 0} \frac{((T_{t+h} - T_t)u)x}{h},$$

and we have the following lemma.

Lemma 5.1. *([18] Lemma 4.1.14) Let $(T_t)_{t\geq 0}$ be a strongly continuous semigroup of operators on a Banach space $(X, \|\cdot\|_X)$, and denote by A its generator with domain $D(A) \subset X$, then*

(i) For any $u \in X$ and $t \geq 0$, it follows that $\int_0^t T_s u ds \in D(A)$ and

$$T_t u - u = A \int_0^t T_s u ds.$$

(ii) For $u \in D(A)$ and $t \geq 0$, we have $T_t u \in D(A)$, that is, $D(A)$ is invariant under T_t, and

$$\frac{d}{dt} T_t u = A T_t u = T_t A u.$$

(iii) For $u \in D(A)$ and $t \geq 0$, we get

$$T_t u - u = \int_0^t A T_s u ds = \int_0^t T_s A u ds.$$

The derivative $\frac{d}{dt} T_t u$ is well defined on the domain $D(A)$ of A and in fact the equation $D_t T_t f = A T_t f$ is Kolmogorov's backward equation and $D_t T_t f = T_t A f$ is Kolmogorov's forward equation.

From a stochastic point of view, operator semigroups start from the study of Markov processes.

Definition 5.3. *Given a Markov process X, we can define the corresponding family of operators $(T_{s,t})$ for $0 \leq s \leq t$ by*

$$(T_{s,t} f)(x) = E[f(X(t)) | X(s) = x], \tag{5.3}$$

for each $f \in B_b(\mathbb{R}^n)$, $x \in \mathbb{R}^n$, where $B_b(\mathbb{R}^n)$ denotes the space of bounded Borel measurable functions on $\mathbb{R}^n$.

If a Markov process $X(t)$ is time-homogeneous, we can write $T_{t-s} = T_{s,t}$.

Theorem 5.1. *Let $X(t)$ be a time-homogeneous Markov process, then the transition operators $(T_t)_{t\geq 0}$ form a semigroup.*

Proof. We want to show that $T_{t+s}f(x) = T_t T_s f(x)$ holds for any $f \in B_b(\mathbb{R}^n)$. We have

$$T_s f(x) = T_{0,s} f(x) = E^x[f(X(s))] = g(x).$$

Then, because of the Markov property of X, we get

$$\begin{aligned} T_t(T_s f(x)) &= T_{0,t} g(x) = E^x[g(X(t))] = E^x[E^{X(t)}[f(X(s))]] \\ &= E^x[E^x[f(X_{s+t})]] = E^x[f(X_{s+t})] = T_{s+t} f(x). \end{aligned}$$

Hence, $T_{s+t} = T_s T_t$, $(T_t)_{t\geq 0}$ form a semigroup. □

Example 5.1. *(The generator of a compound Poisson process)*
Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent and identically-distributed random variables with distribution function $F(x)$ and let N_t be a Poisson process with intensity λ. Denote $S_n = X_1 + \cdots + X_n$. The compound Poisson process Y is defined by

$$Y(t) = \sum_{n\geq 1} S_n 1_{N_t = n}.$$

The transition operator T_t of Y is given by

$$T_t = E^x[f(Y(t))],$$

where we assume that $g \in C_b(\mathbb{R})$. To simplify calculations, we define an operator L by

$$Lf(x) := E[f(x + X_1)] = \int_{\mathbb{R}} f(x+y) F(dy)$$

and note that

$$L^n f(x) := E[f(x + X_1 + \cdots + X_n)] = E[f(x + S_n)].$$

Now it holds that

$$\begin{aligned}(T_t f)(x) = E^x[f(Y(t))] &= \sum_{n\geq 0} E[f(x+S_n)]P(N_t = n) \\ &= \sum_{n\geq 0} e^{-\lambda t}\frac{(\lambda t)^n}{n!}E[f(x+S_n)] \\ &= \sum_{n\geq 0} e^{-\lambda t}\frac{(\lambda t)^n}{n!}L^n f(x) \\ &= \Big(e^{\lambda t(L-I)f} f\Big)(x),\end{aligned}$$

and the transition semigroup can be written as e^{tA} *where* A *is given by*

$$Af(x) = \lambda(L-I)f(x) = \lambda\int_{\mathbb{R}}\Big(f(x+y) - f(x)\Big)F(dy).$$

□

Example 5.2. *The generator of a Levy process)*

Let $(X(t))_{t\geq 0}$ *be a Levy process on* $\mathbb{R}^n$ *with characteristic triple* (A, μ, γ)*. Then the generator of* X *is defined for any* $u \in C_0(\mathbb{R})$

$$\begin{aligned}(Au)(x) =& \frac{1}{2}\sum_{j,k=1}^{n} A_{jk}\frac{\partial^2 u}{\partial x_j x_k}(x) + \sum_{j=1}^{n}\frac{\partial u}{\partial x_j}(x) \\ &= \int_{\mathbb{R}^n}(f(x+y) - f(x) - \sum_{j=1}^{n} y_j\frac{\partial u}{\partial x_j}(x)1_{|y|\leq 1})\mu(dy).\end{aligned}$$

Now we define the Feller process, a type of process that is essentially a Markov process satisfying some additional mild regularity assumptions.

Definition 5.4. *(Feller process [18] Definition 4.1.4) Let* $(T_t)_{t\geq 0}$ *be a strongly continuous semigroup of operators on a Banach space* $(C_\infty(\mathbb{R}^n, \mathbb{R}), \|\cdot\|_\infty)$ *which is positive preserving, i.e.* $u \geq 0$ *yields* $T_t u \geq 0$*. Then* $(T_t)_{t\geq 0}$ *is called a Feller semigroup.*

A Markov process X *with transition semigroup* $(T_t)_{t\geq 0}$ *is a Feller process if* $(T_t)_{t\geq 0}$ *is a Feller semigroup.*

The class of Feller processes includes Lévy processes, Dupire local volatility processes and affine processes in finance. Feller processes may have nonstationary increments, while Lévy processes necessarily have stationary increments.

Recall that the positive maximum principle [12, Thereom 4 in Section 6.4] also holds for an elliptic second order differential operator

$$Lu = \sum_{i,j=1}^{n} a_{ij}(x)u_{x_i x_j} + \sum_{i=1}^{n} b_i(x)u_{x_i} + c(x)u.$$

The connection to semigroups is made by fact that generators of Feller processes satisfy the same maximum principle.

Proposition 5.1. *(Maximum principle [18, Theorem 4.5.1])*
Let $(T_t)_{t\geq 0}$ be a Feller semigroup on $C_\infty(\mathbb{R}^n, \mathbb{R})$ with generator $(A, D(A))$, $D(A) \subset C_\infty(\mathbb{R}^n, \mathbb{R})$. Then A satisfies the positive maximum principle; that is, for $u \in D(A)$ such that for some $x_0 \in \mathbb{R}^n$ the fact that $u(x_0) = \sup_{x\in\mathbb{R}^n} u(x) \geq 0$ implies that $Au(x_0) \leq 0$.

Proof. Suppose that $u \in D(A)$ and that for some $x_0 \in \mathbb{R}^n$ we have $u(x_0) = \sup_{x\in\mathbb{R}^n} u(x) \geq 0$. Since each of the operators T_t, $t \geq 0$ is positivity preserving we find that

$$(T_t u)(x_0) = (T_t u^+)(x_0) - (T_t u^-)(x_0) \leq (T_t u^+)(x_0) \leq \|u^+\|_\infty = u(x_0)$$

which implies

$$Au(x_0) = \lim_{t\mapsto 0} \frac{T_t u(x_0) - u(x_0)}{t} \leq 0.$$

□

The fact that generators of a Feller semigroup and elliptic operators satisfy the positive maximum principle suggests a connection between them. Denote by $C_c^\infty(\mathbb{R}^n)$ the class of functions on $\mathbb{R}^n$ which are infinitely differentiable and have compact support. Then we recall the following theorem.

Theorem 5.2. *[19] If X is a continuous Feller processes on $[0, T]$ with operator A and $C_c^\infty(\mathbb{R}) \subseteq D(A)$, then A is elliptic.*

5.2 Pseudo-differential operators

In the preceding section we have seen that if X is a continuous Feller process, its generator A is a second order elliptic differential operator. As an example, consider an $\mathbb{R}$-valued Lévy process X and its transition operator T_t:

$$T_t f(x) = E[f(X(t))|X(0) = x] = \int_{\mathbb{R}} f(x + y)\mu_t(dy).$$

The second equation is true because of independence and stationarity of increments; μ_t is a probability measure. The Fourier transform of μ_t is

$$\widehat{\mu_t}(u) = \int_{\mathbb{R}} e^{ixu}\mu_t(dx) = E[e^{iuX(t)}|X(0) = x] = e^{t\phi(u)},$$

where ϕ is the characteristic exponent of the Levy process X. Using the convolution theorem, we get

$$\widehat{T_t f}(u) = \widehat{f * \mu_t}(u) = \widehat{f}(u)\widehat{\mu_t}(u) = \widehat{f}(u)e^{t\phi(u)}.$$

The inverse Fourier transform gives

$$T_t f(x) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-iux}e^{t\phi(u)}\widehat{f}(u)du.$$

The generator A of X is given by

$$\begin{aligned} Af(x) &= \lim_{t\to 0}\frac{T_t f(x) - f(x)}{t} \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-iux}\frac{e^{t\phi(u)} - 1}{t}\widehat{f}(u)du \\ &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-iux}\phi(u)\widehat{f}(u)du. \end{aligned}$$

In general, operators with such a representation are called pseudo-differential operators. We give a formal definition, beginning with

Definition 5.5. *(Continuous negative definite functions)*
A function $\phi : \mathbb{R}^n \to C$ is continuous negative definite if it is continuous and if, for any choice of $k \in N$ and vectors $\xi^1, \ldots, \xi^k \in \mathbb{R}^n$, the matrix

$$(\phi(\xi^j) + \overline{\phi(\xi^l)} - \phi(\xi^j - \xi^l))_{j,l=1,\ldots,k}$$

is positive Hermitian, i.e. for all $c_1, \ldots, c_n \in C$,

$$\sum_{j,l=1}^{m}(\phi(\xi^j) + \overline{\phi(\xi^l)} - \phi(\xi^j - \xi^l))_{j,l=1,\ldots,k}c_j\bar{c}_l \geq 0.$$

Some typical examples of continuous negative definite functions are:

- $|\xi|^\alpha$ for $\alpha \in (0, 2]$,
- $1 - e^{-is\xi}$ for $s \geq 0$,

- $\log(1+\xi^2) + i \arctan \xi$.

We now recall

Definition 5.6. *(Pseudo-differential operators) [18]*
Let $(A, D(A))$ be an operator with $C_0^\infty(\mathbb{R}^n) \subset D(A)$. Then A is a pseudo-differential operator if

$$\begin{aligned}(Au)(x) &= -a(x,D)u(x) \\ &= -(2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{ix\cdot\xi} a(x,\xi)\hat{u}(\xi) d\xi,\end{aligned} \tag{5.4}$$

for $u \in C_0^\infty(\mathbb{R}^n)$.

$$\hat{u}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(x) dx$$

is the Fourier transform of u. The symbol $a(x,\xi) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is locally bounded in (x,ξ), $a(\cdot,\xi)$ is measurable for every ξ, and $a(x,\cdot)$ is a continuous negative definite function for every x.

Time-inhomogeneous processes

Definition 5.7. *The family of generators of X is defined by*

$$A_s u = \lim_{h\to 0^+} \frac{T_{s-h,s}u - u}{h}, \tag{5.5}$$

for all $s > 0$, $f \in D(A_s)$, where $D(A_s)$ denotes the domain of A_s.

Definition 5.8. *Let $(A_s)_{s>0}$ be a family of operators with $C_0^\infty(\mathbb{R}^n) \subset D(A_s)$. Then A_s is a pseudo-differential operator for all $s > 0$ if*

$$\begin{aligned}(A_s u)(x) &= -a(s,x,D)u(x) \\ &= -(2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{ix\cdot\xi} a(s,x,\xi)\hat{u}(\xi) d\xi,\end{aligned} \tag{5.6}$$

for $u \in C_0^\infty(\mathbb{R}^n)$, $\hat{u}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(x) dx$, is the Fourier transform of u. The symbol $a(s,x,\xi) : \mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^n \to C$ is locally bounded in (x,ξ), $a(s,\cdot,\xi)$ is measurable for every ξ, s, and $a(s,x,\cdot)$ is a continuous negative definite function for every (s,x).

5.3 Construction of a Markov process using a symbol

Time-homogeneous case

In [17], Hoh showed that if a symbol of a pseudo-differential operator satisfies certain conditions, then one can construct a unique Markov process whose generator has that symbol.

Let (E, d) be a separable metric spaceand let D_E denote the space of all cadlag paths with values in E,

$$D_E := \{\omega : [0, \infty) \to E, \omega \text{ is right continuous, } \lim_{s \to t} \omega(s) \text{ exists for all } t > 0\},$$

let $M(D_E)$ denote the set of probability measures on D_E.

Definition 5.9. *[17] Let A be a linear operator with domain $D(A)$. A probability measure $P \in M(D_E)$ is called a solution of the martingale problem for the operator A if for every $\phi \in D(A)$, the process*

$$\phi(X(t)) - \int_0^t A\phi(X(s))ds$$

is a martingale under $\mathbb{P}$ with respect to the filtration $\mathbb{F} = \{\mathscr{F}()_{t \geq 0}$.
If for every probability measure $\mu \in M(D_E)$, there is a unique solution P_μ of the martingale problem for A with initial distribution

$$P_\mu \circ X(0)^{-1} = \mu,$$

then the martingale problem for A is called well posed.

We assume that $\phi : \mathbb{R}^n \to \mathbb{R}$ is a continuous negative definite reference function for some $r > 0$, and $c > 0$. Define $\lambda(\xi) = (1 + \phi(\xi))^{1/2}$, $\xi \in \mathbb{R}^n$.

Theorem 5.3. *([17])*
Let $a : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a continuous negative definite symbol such that $a(x, 0) = 0$ for all $x \in \mathbb{R}^n$. Let M be the smallest integer such that

$$M > (\frac{n}{r} \vee 2) + n$$

and suppose that

(i) $a(x,\xi)$ *is* $(2M+1-n)$ *times continuous differentiable with respect to* x *and for all* $\beta \in \mathbb{N}$, $|\beta| \leq 2M+1-n$,

$$|\partial_x^\beta a(x,\xi)| \leq c\lambda^2(\xi) \tag{5.7}$$

holds for all $x \in \mathbb{R}^n$, $\xi \in \mathbb{R}^n$.

(ii) For some strictly positive function $\gamma : \mathbb{R}^n \to \mathbb{R}^+$,

$$a(x,\xi) \geq \gamma(x) \cdot \lambda^2(\xi), \tag{5.8}$$

for all $x \in \mathbb{R}^n$, $|\xi| \geq 1$. *Then the martingale problem for the operator* $-a(x,D)$ *with domain* $C_0^\infty(\mathbb{R}^n)$ *is well-posed.*

We illustrate theorem 5.3 with some examples. Let the continuous negative definite reference function $\phi(\xi)$ be ξ^2, then $\lambda(\xi) = (1+\xi^2)^{1/2}$.

Examples:

1. Brownian motion $W(t)$, the symbol of its generator is

$$a(x,\xi) = \frac{1}{2}\xi^2.$$

2. Geometric Brownian motion $X(t)$, $dX(t) = \mu X(t)dt + \sigma X(t)dW(t)$, the symbol of its generator is

$$a(x,\xi) = \frac{1}{2}\sigma^2 x^2 \xi^2 - \mu x \xi.$$

3. CIR process $X(t)$, $dX(t) = \theta(\mu - X(t))dt - \sigma\sqrt{X(t)}dW(t)$, the symbol of its generator is

$$a(x,\xi) = \frac{1}{2}\sigma^2 x \xi^2 + \theta(\mu - x)\xi.$$

4. Levy process $X(t)$ with Levy triplet (σ,μ,γ), the symbol of its generator is

$$a(x,\xi) = c(x) - i\gamma\xi + \frac{1}{2}\sigma^2\xi^2 + \int_{\mathbb{R}} (1 - e^{iy\xi} + \frac{iy\xi}{1+y^2})\nu(dy).$$

5. A continuous diffusion process $X(t)$, $dX(t) = \sigma(X(t))dW(t)$ where $\sigma(X(t))$ ia an adapted process, the symbol of its generator is

$$a(x,\xi) = \frac{1}{2}\sigma(x)^2\xi^2.$$

In order to satisfy (5.7), when $M = 3$, $n = 1$, $\sigma(x)$ needs to be 6 times continuous differentiable with respect to x. If $\gamma(x) = \frac{1}{4}\sigma(x)^2$, then (5.8) holds for all $|\xi| \geq 1$.

Time-inhomogeneous case

Bottcher showed that we could construct a time-inhomogeneous Markov process using pseudo-differential operators.

Definition 5.10. *([6]) A continuous negative definite function $\phi : \mathbb{R}^n \to \mathbb{R}$ belongs to the class Λ if for all $\alpha \in N_0^n$ there exists constants $c_\alpha \geq 0$ such that*

$$|\partial_\xi^\alpha (1 + \phi(\xi))| \leq c_\alpha (1 + \phi(\xi))^{2-(|\alpha \wedge 2|)/2}.$$

Definition 5.11. *([6]) Let $m \in \mathbb{N}$, $j \in 0, 1, 2$ and $\phi \in \Lambda$. A function $a : \mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is in the class $S_j^{\phi,m}$ if for all $\alpha, \beta \in N_0^n$ and for any compact $K \subset \mathbb{R}^+$ there are constants $c_{\alpha,\beta,K} \geq 0$ such that*

$$|\partial_\xi^\alpha \partial_x^\beta a(t, x, \xi)| \leq c_{\alpha,\beta,K} (1 + \phi(\xi))^{m-(|\alpha| \wedge j)}/2$$

holds for all $t \in K$, $x \in \mathbb{R}^n$ and $\xi \in \mathbb{R}^n$. Here $m \in \mathbb{R}$ is called the order of the symbol. Furthermore, the notation $a \in (t - s)S_j^{\phi,m}$ is used if, for $s \leq t$,

$$|\partial_\xi^\alpha \partial_x^\beta a(s, t, x, \xi)| \leq (t - s)c_{\alpha,\beta} (1 + \phi(\xi))^{m-(|\alpha| \wedge j)}/2$$

holds, where the constants $c_{\alpha,\beta}$ are independent of s and t.

We now recall

Theorem 5.4. *([6, Theorem 4.2])*
Suppose $\phi \in \Lambda$ is a negative definite function, and there exist $c_0 > 0$, $r_0 > 0$ such that $\phi(\xi) \geq c_0 |\xi|^{r_0}$ for all ξ large. If a pseudo-differential operator with symbol $a(s, x, \xi)$ satisfies the following conditions,

- *$a(\cdot, x, \xi)$ is a continuous function for all $x \in \mathbb{R}^n$, $\xi \in \mathbb{R}^n$,*
- *$a(t, x, \cdot)$ is continuous negative definite for all $t \in \mathbb{R}^+$, $x \in \mathbb{R}^n$,*
- *$a(t, x, 0) = 0$ holds for all t and x,*

- *$a \in S_2^{\phi,m}$ is elliptic, that is, there exist $R > 0$, $c > 0$, for any x, $|\xi| \geq R$, $\Re a(t,x,\xi) \geq c(1+\phi(\xi))^{m/2}$ holds uniformly in t on compact sets,*

then $a(s,x,\xi)$ defines a family of operators $T_{s,t}$ on C^∞ such that

(a) $T_{s,t}$ is a linear operator,

(b) $T_{s,t}$ is a contraction,

(c) $T_{s,t}$ is positivity preserving,

(d) $T_{t,t} = I$,

(e) $T_{s,t} = T_{s,\tau}T_{\tau,t}$ for $s \leq \tau \leq t$,

(f) $T_{s,t}1 = 1$, that is, $\lim_{k\to\infty} T_{s,t}u_k = 1$ holds for $u_k \in C_\infty$ with $u_k \to 1$.

Theorem 5.4 has the following important corollary:

Corollary 5.1. *([6, Corallary 4.3]) The operator given in Theorem 5.4 defines a Markov process.*

Example:

1. Dupire local volatility model $dS_t = \mu S_t\, dt + \sigma(t,S_t)S_t\, dW(t)$

$$a(t,x,\xi) = \frac{1}{2}\sigma(t,x)^2x^2\xi^2 - \mu x\xi$$

2. A time-inhomogeneous Markov process in $\mathbb{R}^n$ with generator

$$\begin{aligned} A_t u(t,x) = &\sum_{i,j} a_{ij}(t,x)\frac{\partial^2 u(t,x)}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(t,x)\frac{\partial u(t,x)}{\partial x_i} + c(t,x)u(t,x) \\ &+ \int_{\mathbb{R}^n\setminus\{0\}} \big(u(t,x+y) - u(t,x) - \frac{< y, \nabla u(t,x) >}{1+|y|^2}\big)\mu(t,x,dy) \end{aligned}$$

$$\begin{aligned} a(t,x,\xi) = &\frac{1}{2}\xi^T(a_{ij(t,x)})_{ij}\xi - b(t,x)\cdot\xi + c(t,x) \\ &+ \int_{\mathbb{R}^n\setminus\{0\}} \big(1 - e^{iy\cdot\xi} + \frac{iy\cdot\xi}{1+|y|^2}\big)\mu(t,x,dy). \end{aligned}$$

Future research plan

Suppose there is a semimartingale $X(t)$ with decomposition (4.2)

$$X(t) = X(0)+\int_0^t \beta(s)ds+\int_0^t \sigma(s)dW(s)+\int_0^t \int_{\|y\|\leq 1} y\widetilde{M}(dsdy)+\int_0^t \int_{\|y\|>1} yM(dsdy).$$

In section 4.2, we derive the generator A of X,

$$\begin{aligned} Af(x) =& \sum_{i=1}^n b_i(x,t)\frac{\partial f}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x,t)\frac{\partial^2 f}{\partial x_i x_j} \\ &+ \int_{\mathbb{R}^n} \big(f(x+y) - f(x) - 1_{\|y\|\leq 1}\sum_{i=1}^n y_i \frac{\partial f}{\partial x_i}\big)\nu(t,y,x)dy. \end{aligned}$$

We plan to show that A is a pseudodifferential operator with symbol $a(t, x, \xi)$ as defined in Definition 5.8. We would like to discover what are the requirements for the coefficients a_{ij}, b_i, c, n in the generator A, such that $a(t, x, \xi)$ satisfies the conditions in Theorem 5.4. If we had this result, we could conclude that there exists a unique Markov process $Y(t)$ which mimics $X(t)$ for all $0 \leq t \leq T$.

Chapter 6

Application of Brunick's Theorem to Barrier Options

6.1 Introduction

In practice, the local volatility model is used to price both vanilla options with path-independent payoffs and complex options with path-dependent payoffs. However, the price of an option with a path-dependent payoff cannot be uniquely determined by the marginal distributions of the asset price process. For example, the price of a barrier option would require knowledge of the joint probability distribution of the asset price process and its running maximum. Brunick [8] generalizes Gyöngy's result under a weaker assumption so that the prices of path-dependent options can be determined exactly.

Definition 6.1. *([8, Definition 2.1])*
Let $Y : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^n$ be a predictable process. We say that the path-functional Y is a ***measurably updatable statistic*** *if there exists a measurable function*

$$\phi : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^n$$

such that $Y(t + u, x) = \phi(t, Y(t, x); u, \Delta(t, x))$ for all $x \in C(\mathbb{R}_+, \mathbb{R}^d)$, where the map $\Delta : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^D) \to C(\mathbb{R}_+, \mathbb{R}^d)$ is defined by $\Delta(t, x)(u) = x(t, u) - x(t)$.

A measurably updatable statistic is a functional whose path-dependence can be summed up by a single vector in $\mathbb{R}^N$. $Y^1(t, x) = x(t)$ is an updatable statistic, as $Y^1(t + u, x) = Y^1(t, x) + \Delta(t, x)(u)$. Let $x^*(t) = \sup_{u \le t} x(u)$, we see that $Y^2(t, x) = [x(t), x^*(t)] \in \mathbb{R}^2$ is an updatable statistic as we can write

$$Y^2(t + u, x) = [x(t) + \Delta(t, x)(u), \max\{x^*(t), \sup_{v \in [0,u]} x(t) + \Delta(t, x)(v)\}].$$

We first recall Brunick's extension of Gy¨ngy's theorem.

Theorem 6.1. *([8, Theorem 2.11])*
Let W be an r-dimensional Brownian motion, and

$$dX(t) = \mu_t dt + \sigma(t)dW(t) \tag{6.1}$$

be a d-dimensional Itô process where μ is a left-continuous d-dimensional adapted process, and σ is a left-continuous $d \times r$-dimensional adapted process with

$$E[\int_0^t |\mu_s| + |\sigma(s)\sigma(s)^T ds|ds] \leq \infty \text{ for all } t.$$

Also suppose that $Y : \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d) \to \mathbb{R}^n$ is a measurably updatable statistic such that the maps $x \mapsto Y(t,x)$ are continuous for each fixed t. Then there exist deterministic measurable functions $\hat{\mu}$ and $\hat{\sigma}$ such that

$$\begin{aligned} \hat{\mu}(t, Y(t,X)) &= E[\mu_t | Y(t,X)] \text{ a.s. for Lebesgue-a.e. } t, \\ \hat{\sigma}\hat{\sigma}^T(t, Y(t,X)) &= E[\sigma(t)\sigma(t)^T | Y(t,X)] \text{ a.s. for Lebesgue-a.e. } t, \end{aligned}$$

and there exists a weak solution to the SDE:

$$d\hat{X}_t = \hat{\mu}(t, Y(t,\hat{X}))dt + \hat{\sigma}(t, Y(t,\hat{X}))d\hat{W}(t)$$

such that $\mathscr{L}(Y(t,\hat{X})) = \mathscr{L}(Y(t,X))$ for all $t \in \mathbb{R}_+$, where $\mathscr{L}$ denotes the law of a random variable and $\hat{W}(t)$ denotes another Brownian motion.

Brunick's theorem is more general than Gyöngy's. First, the requirements of μ and σ are weaker than the boundedness and uniform ellipticity in Gyöngy's theorem. Secondly, this theorem gives the existence of a weak solution which preserves the one-dimensional marginal distribution of *path-dependent* functional.

6.2 Application to up-and-out calls

We can apply Brunick's theorem to construct a 2-dimensional Markov process explicitly which mimicks the joint marginal density of an Itô process and its running maximum (or minimum).

For example, consider an up-and-out European call. The payoff of this option is

$$(S(T) - K)^+ 1_{\{S^*(T) \leq B\}},$$

where $S^*(T) = \sup_{0\leq u\leq T} S(u)$ is the running maximum of $S(t)$. The stock price, $S(t)$, follows the SDE

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

where $\sigma(t)$ is any adapted stochastic process, where r is the constant interest rate and $\{W(t)\}_{t\geq 0}$ is Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$. Our goal is to find an analytic formula for

$$\hat{\sigma}(t, x, y) := E[\sigma(t)|S(t) = x, S^*(t) = y].$$

We apply this result to the problem of pricing single barrier options. Theorem ( ) tells us that there exists a measurable, deterministic function $\hat{\sigma}$ such that

$$\hat{\sigma}(t, x, y) = E[\sigma(t)|S(t) = x, S^*(t) = y],$$

and there exists a weak solution to the SDE

$$d\hat{S}(t) = r\hat{S}(t)dt + \hat{\sigma}(t, \hat{S}(t), \hat{S}^*(t))\hat{S}(t)d\hat{W}(t),$$

where $(\hat{S}(t), \hat{S}^*(t))$ has the same joint distribution as $(S(t), S^*_t)$.

The risk-neutral pricing formula tells us that the price of an up-and-out call option with strike K, maturity T and barrier B is given by

$$\begin{aligned} C(T, K, B) &= e^{-rT} \int_{\mathbb{R}} \int_{\mathbb{R}} (S(T) - K)^+ 1_{\{S^*(T)\leq B\}} \phi(T, S(T), S^*(T)) dS(T) dS^*(T) \\ &= e^{-rT} \int_{K}^{\infty} \int_{-\infty}^{B} (S(T) - K) \phi(T, S(T), S^*(T)) dS(T) dS^*(T), \end{aligned}$$

where $\phi(T, S(T), S^*(T))$ is the joint distribution of $(S(T), S^*(T))$. We differentiate both sides with respect to B and get

$$\frac{\partial}{\partial B} C(T, K, B) = e^{-rT} \int_{K}^{\infty} (S(T) - K) \phi(T, S(T), B) dS(T). \tag{6.2}$$

We differentiate both sides with respect to K and get

$$\frac{\partial^2}{\partial B \partial K} C(T, K, B) = -e^{-rT} \int_{K}^{+\infty} \phi(T, S(T), B) dS(T). \tag{6.3}$$

Again, we can differentiate both sides with respect to K and get

$$\frac{\partial^3}{\partial B \partial K^2} C(T,K,B) = -e^{-rT}\phi(T,K,B). \tag{6.4}$$

Hence, the joint transition density function can be represented as a derivative of the up-and-out call option price.

Next we apply the Itô-Tanaka's formula [24, Theorem 43.3] to the function $(\hat{S}(t) - K)^+ 1_{\{\hat{S}^*(T)\le B\}}$, we get

$$\begin{aligned}(\hat{S}(T)-K)^+1_{\{\hat{S}^*(T)\le B\}} &= [(\hat{S}_0-K)^+ + \int_0^T 1_{[K,\infty)}(\hat{S}_u)d\hat{S}_u + \frac{1}{2}L_T^K]1_{\{\hat{S}^*(T)\le B\}} &(6.5)\\ &= [(\hat{S}_0-K)^+ + \int_0^T 1_{[K,\infty)}(\hat{S}_u)\hat{S}_u(rdu &(6.6)\\ &\quad +\hat{\sigma}(u,\hat{S}_u,\hat{S}_u^*)dW(u)) + \frac{1}{2}L_T^K]1_{\{\hat{S}^*(T)\le B\}}\end{aligned}$$

where L_T^K is the local time of

$hatS(T)$ at K. By the definition of local time,

$$\begin{aligned}L_T^K &= \lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\int_0^T 1_{(K-\varepsilon,K+\varepsilon)}(\hat{S}_u)d<\hat{S},\hat{S}>_u\\ &= \lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\int_0^T 1_{(K-\varepsilon,K+\varepsilon)}(\hat{S}_u)\hat{\sigma}^2(u,\hat{S}_u,\hat{S}_u^*)\hat{S}_u^2du.\end{aligned}$$

Taking expectations of (6.5) we get

$$\begin{aligned}C(T,K,B) &= e^{-rT}E[(S(T)-K)^+1_{\{S^*(T)\le B\}}]\\ &= e^{-rT}E[(\hat{S}(T)-K)^+1_{\{\hat{S}(T)^*\le B\}}]\\ &= e^{-rT}\{(\hat{S}_0-K)^+E[1_{\{\hat{S}(T)^*\le B\}}] + E[\int_0^T r1_{[K,\infty)}(\hat{S}_u)\hat{S}_u1_{\{\hat{S}(T)^*\le B\}}du]\\ &\quad +\frac{1}{2}E[\lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\int_0^T 1_{(K-\varepsilon,K+\varepsilon)}(\hat{S}_u)\hat{\sigma}^2(u,\hat{S}_u,\hat{S}_u^*)\hat{S}_u^2du1_{\{\hat{S}(T)^*\le B\}}]\}\\ &= e^{-rT}\{(\hat{S}_0-K)^+E[1_{\{\hat{S}(T)^*\le B\}}] + r\int_0^B\int_K^\infty\int_0^T \hat{S}_u\phi(u,\hat{S}_u,\hat{S}_u^*)dud\hat{S}_ud\hat{S}_u^*\\ &\quad +\frac{1}{2}\int_0^\infty\int_0^B[\lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\int_0^T 1_{(K-\varepsilon,K+\varepsilon)}(\hat{S}_u)\hat{\sigma}^2(u,\hat{S}_u,\hat{S}_u^*)\hat{S}_u^2du]\\ &\quad \phi(u,\hat{S}_u,\hat{S}_u^*)d\hat{S}_ud\hat{S}_u^*\}\\ &= e^{-rT}\{(\hat{S}_0-K)^+P(\hat{S}(T)^*\le B) + r\int_0^B\int_K^\infty\int_0^T \hat{S}_u\phi(u,\hat{S}_u,\hat{S}_u^*)dud\hat{S}_ud\hat{S}_u^*\\ &\quad +\frac{1}{2}\int_0^B\int_0^T \hat{\sigma}^2(u,K,\hat{S}_u^*)K^2\phi(u,K,\hat{S}_u^*)]dud\hat{S}_u^*\},\end{aligned}$$

where $\phi(u, \hat{S}_u, \hat{S}_u^*)$ is the joint density of $(\hat{S}_u, \hat{S}_u^*)$.

Taking derivatives with respect to B gives

$$\begin{aligned}\frac{\partial}{\partial B}C(T,K,B) &= e^{-rT}\Big\{(\hat{S}_0-K)^+\phi_{\hat{S}(T)^*}(B) + r\int_K^\infty\int_0^T \hat{S}_u\phi(u,\hat{S}_u,B)dud\hat{S}_u \\ &\quad +\frac{1}{2}\int_0^T \hat{\sigma}^2(u,K,B)K^2\phi(u,K,B)du\Big\},\end{aligned}$$

and by taking derivatives with respect to T we obtain

$$\begin{aligned}\frac{\partial^2}{\partial B\partial T}C(T,K,B) &= -r\frac{\partial}{\partial B}C(T,K,B) + e^{-rT}\Big\{(\hat{S}_0-K)^+\frac{\partial}{\partial T}\phi_{\hat{S}(t)^*}(B) \\ &\quad +r\int_K^\infty \hat{S}(T)\phi(T,\hat{S}(T),B)d\hat{S}(T) + \frac{1}{2}\hat{\sigma}^2(T,K,B)K^2\phi(T,K,B)\Big\} \\ &= -r\frac{\partial}{\partial B}C(T,K,B) + e^{-rT}\Big\{(\hat{S}_0-K)^+\frac{\partial}{\partial T}\phi_{\hat{S}(t)^*}(B) \\ &\quad +\frac{1}{2}\hat{\sigma}^2(T,K,B)K^2\phi(T,K,B) + r\int_K^\infty K\phi(T,\hat{S}(T),B)d\hat{S}(T) \\ &\quad +r\int_K^\infty (\hat{S}(T)-K)\phi(T,\hat{S}(T),B)d\hat{S}(T)\Big\},\end{aligned}$$

where $\phi_{\hat{S}^*(t)}(B)$ is the marginal distribution of $\hat{S}^*(t)$. Combining (6.2), (6.3) and (6.4) yields

$$\begin{aligned}&rK\frac{\partial^2}{\partial B\partial K}C(T,K,B) + \frac{\partial^2}{\partial B\partial T}C(T,K,B) \\ &= e^{-rT}(\hat{S}_0-K)^+\frac{\partial}{\partial T}\phi_{\hat{S}(T)^*}(B) - \frac{1}{2}\hat{\sigma}^2(T,K,B)K^2\frac{\partial^3}{\partial B\partial K^2}C(T,K,B),\end{aligned}$$

where $\hat{\sigma}(T,K,B)$ is the local volatility of up-and-out call, $C(T,K,B)$ is the price of an up-and-out call.

Consider the density $\phi_{\hat{S}(T)^*}(B)$:

$$\begin{aligned}\phi_{\hat{S}(T)^*}(B) &= \int_0^\infty \phi(T,\hat{S}(T),B)dS(T) \\ &= \int_0^K \phi(T,\hat{S}(T),B)dS(T) + \int_K^\infty \phi(T,\hat{S}(T),B)dS(T).\end{aligned}$$

From (6.3), we have

$$\int_K^{+\infty} \phi(T,S(T),B)dS(T) = -e^{rT}\frac{\partial^2}{\partial B\partial K}C(T,K,B).$$

We can derive a similar formula for $\int_0^K \phi(T,\hat{S}(T),B)dS(T)$ by differentiating the up-and-out put option price.

The up-and-out put with strike K, maturity T and barrier B is given by

$$\begin{aligned} P(T,K,B) &= e^{-rT}\int_{\mathbb{R}}\int_{\mathbb{R}}(K-S(T))^{+}1_{\{S^*(T)\leq B\}}\phi(T,S(T),S^*(T))dS(T)dS_T^* \quad (6.7)\\ &= e^{-rT}\int_0^K\int_{-\infty}^B(K-S(T))\phi(T,S(T),S^*(T))dS(T)dS_T^*. \end{aligned}$$

We differentiate (6.6) with respect to B to get

$$\frac{\partial}{\partial B}P(T,K,B) = e^{-rT}\int_0^K(K-S(T))\phi(T,S(T),B)dS(T).$$

Differentiating the preceding expression with respect to K yields

$$\frac{\partial^2}{\partial B\partial K}P(T,K,B) = e^{-rT}\int_0^K\phi(T,S(T),B)dS(T).$$

Therefore,

$$\phi_{\hat{S}^*(T)}(B) = e^{rT}\left\{\frac{\partial^2}{\partial B\partial K}P(T,K,B) - \frac{\partial^2}{\partial B\partial K}C(T,K,B)\right\},$$

and

$$\frac{\partial}{\partial T}\phi_{\hat{S}^*(T)}(B) = r\phi_{\hat{S}(T)^*}(B) + e^{rT}\left\{\frac{\partial^3}{\partial B\partial K\partial T}P(T,K,B) - \frac{\partial^3}{\partial B\partial K\partial T}C(T,K,B)\right\}.$$

Substituting the preceding expression into (6.7), we obtain

$$\begin{aligned} &rK\frac{\partial^2}{\partial B\partial K}C(T,K,B) + \frac{\partial^2}{\partial B\partial T}C(T,K,B)\\ &= (\hat{S}_0-K)^+\left\{\frac{\partial^3}{\partial B\partial K\partial T}P(T,K,B) - \frac{\partial^3}{\partial B\partial K\partial T}C(T,K,B) + e^{-rT}r\phi_{\hat{S}(T)^*}(B)\right\}\\ &-\frac{1}{2}\hat{\sigma}^2(T,K,B)K^2\frac{\partial^3}{\partial B\partial K^2}C(T,K,B), \end{aligned}$$

Solving this equation for $\hat{\sigma}(T,K,B)$ yields

$$\begin{aligned} \hat{\sigma}^2(T,K,B) &= \frac{1}{\frac{1}{2}K^2\frac{\partial^3}{\partial B\partial K^2}C}\left\{(\hat{S}_0-K)^+\left[\left(\frac{\partial^3}{\partial B\partial K\partial T}P - \frac{\partial^3}{\partial B\partial K\partial T}C\right) + r\left(\frac{\partial^2}{\partial B\partial K}P\right.\right.\right.\\ &\left.\left.\left.-\frac{\partial^2}{\partial B\partial K}C\right)\right] - rK\frac{\partial^2}{\partial B\partial K}C - \frac{\partial^2}{\partial B\partial T}C\right\}. \end{aligned}$$

This is an analogue of Dupire local volatility for barrier option.

References

[1] M. Atlan. Localizing volatilities. *arXiv:math/0604316v1*, 2006.

[2] D. Bates. Jumps and stochastic volatility: the exchange rate processes impliciti in deutschemark options. *Rev. Fin. Studies,*, 9:69–107, 1996.

[3] A. Bentata and R. Cont. Forward equations for option prices in semimartingale models. *arXiv:1001.1380v2*.

[4] A. Bentata and R. Cont. Mimicking the marginal distributions of a semimartingale. *arXiv:0910.3992v2*.

[5] F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.

[6] B. Böttcher. Construction of time-inhomogeneous Markov processes via evolution equations using pseudo-differential operators. *J. Lond. Math. Soc. (2)*, 78(3):605–621, 2008.

[7] D.T. Breeden and R.H. Litzenberger. Prices of state-contingent claims implicit in option prices. *The Journal of Business*, 51(4):621–651, 1978.

[8] G. Brunick. A weak existence result with application to the financial engineer's calibration problem. *Ph.D dissertation, Carmegie Mellon University*, 2008.

[9] E. Derman and I. Kani. Riding on a smile. *Risk*, 7(2):32–39, 1994.

[10] B. Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[11] B. Dupire. A unified theory of volatility. *Discussion paper Paribas Capital Markets, reprinted in "Derivative Pricing: The Classic Colletion",edited by P. Carr, 2004, Risk Books, London*, 1996.

[12] L.C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[13] P. Feehan. Notes on existence and uniqueness for parabolic differential equations on euclidean space.

[14] A. Friedman. *Partial differential equations of parabolic type*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.

[15] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probab. Theory Relat. Fields*, 71(4):501–516, 1986.

[16] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

[17] W. Hoh. A symbolic calculus for pseudo-differential operators generating Feller semigroups. *Osaka J. Math.*, 35(4):789–820, 1998.

[18] N. Jacob. *Pseudo differential operators and Markov processes. Vol. I Fourier analysis and semigroups.* Imperial College Press, London, 2001.

[19] N. Jacob. *Pseudo differential operators & Markov processes. Vol. II Generators and their potential theory.* Imperial College Press, London, 2002.

[20] N.V. Krylov. On the relation between differential operators of second order and the solutions of stochastic differential equations. *Steklov Seminar. Publication Division, New York*, pages 214–229, 1984.

[21] R. Merton. Option pricing when underlying stock returns are discontinuous,. *J. Financial Ecomonics*, 3:125–144, 1976.

[22] A. Lesniewski P. Hagan, D. Kumar and D. Woodward. Managing smile risk. *Wilmott Magazine*, pages 84–108, 2002.

[23] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[24] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[25] M. Shi. Local intensity and its dynamics in multi-name credit derivatives modeling. *Ph.D. dissertation, Rutgers University*, 2010.

[26] R. E. Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*, volume 49 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 1997.

[27] D.W. Stroock. Diffusion processes associated with Lévy generators. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 32(3):209–244, 1975.

Vita

Jin Wang

2004-2010 Ph. D. in Mathematics, Rutgers, the University of New Jersey, New Brunswick, NJ, USA

1999-2003 B. Sc. in Mathematics from Tsinghua University, Beijing, China

2005-2010 Teaching assistant, Department of Mathematics, Rutgers, the University of New Jersey, New Brunswick, NJ, USA