SOLVABLE MARKOV PROCESSES

by

Alexey Kuznetsov

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Mathematics
University of Toronto

Abstract

Solvable Markov Processes

Alexey Kuznetsov

Doctor of Philosophy

Graduate Department of Mathematics

University of Toronto

2004

This thesis introduces a new method of constructing analytically tractable (solvable) one-dimensional Markov processes - processes for which the transition probability density can be computed explicitly. For this purpose we introduce the concept of stochastic transformation as a way to map solvable diffusion process into a solvable driftless process. Stochastic transformations consist of an absolutely continuous measure change and a diffeomorphism. Our main theorem characterizes all stochastic transformations and gives a remarkably simple algorithm to construct all the stochastic transformations for a given process.

We study in detail properties of these transformations and the boundary behavior of transformed processes. As examples we show how one can obtain the well known and widely used martingale models (quadratic volatility, CEV processes) as well as obtain new important classes (Ornstein-Uhlenbeck, CIR and Jacobi families) of solvable driftless processes.

Inspired by these ideas, in the next chapter we classify all solvable one-dimensional driftless diffusions, for which the transition probability function can be computed as an integral over hypergeometric functions.

In the last chapter we give several examples of how one could use the above processes for financial modelling. First we prove the result that lattice approximations to Ornstein-Uhlenbeck and CIR processes, given by Charlier and Meixner processes, are affine and compute explicitly the generating and characteristic functions of these processes. Then using the eigenfunction expansion of the probability semigroup and the concept of stochastic time change we show how to introduce stochastic volatility and jumps while preserving solvability of our model. As an example of possible applications of these processes we find an explicit formula for the price of a call option and give an explicit algorithm for pricing American style options.

Acknowledgements

I would like to thank:

my supervisor, Professor Claudio Albanese, who inspired my interest in this subject and from whom I learned a great variety of methods and ideas,

Professors Ken Jackson and Jeremy Quastel for sharing their time and ideas,

our graduate coordinator, Ida Bulat, for all her help and support,

my colleagues Oliver Chen and Stephan Lawi, for being such great friends and always being very helpful,

my high school teacher Nahim Borisovich Levit, who was the first person to inspire my love for the beauty and harmony of Mathematics,

the Department of Mathematics, University of Toronto, Ontario Graduate Scholarship programme and the National Sciences and Engineering Research Council of Canada for financial support,

and most of all, my dear family for their love and support.

Contents

Notations

- $X_t = (X_t, \mathrm{P})$ – One-dimensional stationary Markov process under the measure P
- P_x – probability measure associated to X_t started at x
- $D = D_x = [D_x^1, D_x^2] \subseteq \mathbb{R}$ – domain of the process X_t
- $\zeta = \zeta_X = \{\inf s: \ X_s \notin D_x\}$ – the lifetime of the process X_t
- $P_X(t)$ – Markov semigroup associated with the process (X_t, P).
- $\mathcal{L}_X$ – Markov generator of the process X_t (or of the semigroup $P_X(t)$)
- $b_X(x)$ and $\sigma_X(x)$ – drift and volatility coefficients (in case X_t is a diffusion)
- $m_X(dx)$, $s_X(x)$ and $k_X(dx)$ – speed measure, scale function and killing measure of the process (X_t, P)
- $\{\lambda_n\}_{n\geq 0}$ and $\{\psi_n(x)\}_{n\geq 0}$ - spectrum and eigenfunctions of $\mathcal{L}_X$ in the space $L^2(D_x, m_X(dx))$
- $p_X(t, x_0, x_1)$ – transitional probability density of the process X_t with respect to the speed measure $m_X(dx_1)$
- $G_X(\lambda, x_0, x_1)$ – Green function of the process X_t
- $H_z = \{\inf s: \ X_s = z\}$ – the first hitting time of $z \in D_x$
- $h(x)$ – ρ-excessive function: $e^{-\rho t} P_X(t) h(x) \leq h(x)$
- P^h – h-transform of the measure P: $d\mathrm{P}_t^h = e^{-\rho t} h(X_t) d\mathrm{P}_t$
- $\varphi_\rho^+(x)$ and $\varphi_\rho^-(x)$ – increasing and decreasing solutions of $\mathcal{L}_X \varphi(x) = \rho\varphi(x)$

- X_t^h – h-transform (ρ-excessive transform) of the process X_t (process X_t under the measure P^h)

- $p_{X^h}(t, x_0, x_1)$, $G_{X^h}(\lambda, x_0, x_1)$, $\mathcal{L}_{X^h}$

  $b_{X^h}(x)$ and $\sigma_{X^h}(x)$,

  $m_{X^h}(dx)$, $s_{X^h}(x)$ and $k_{X^h}(dx)$ - characteristics of the process $X_t^h = (X_t, \mathrm{P}^h)$.

- $Y(x)$ –diffeomorphism: $Y(x): \ D_x \to D_y \subseteq \mathbb{R}$ and $X(y) = Y^{-1}(y)$ - its inverse ($x_0 = X(y_0)$ and $x_1 = X(Y_1)$)

- $\{\rho, h, Y\}$ – stochastic transformation

- $Y_t = Y(X_t^h)$ – driftless process $Y_t = (Y(X_t), \mathrm{P}^h)$ on the domain $D_y = Y(D_x)$

- $p_Y(t, y_0, y_1)$, $G_Y(\lambda, y_0, y_1)$, $\mathcal{L}_Y$, $\sigma_Y(y)$,

  $m_Y(dy)$, $s_Y(y)$ and $k_Y(dy)$ - characteristics of the process $Y_t = (Y(X_t), \mathrm{P}^h)$.

- $C_Y(t, y_0, K)$ – price of a European call option $E^{\mathrm{P}^h}((Y_t - K)^+ | Y_0 = y_0)$

- $(X, \mathrm{P}) \sim (Y, \mathrm{Q})$ – (X, P) and (Y, Q) are related by a stochastic transformation

- $\mathfrak{M}(X, \mathrm{P}) = \{(Y, \mathrm{Q}) : \ (Y, \mathrm{Q}) \sim (X, \mathrm{P})\}$ – class of driftless processes related to (X, P) by a stochastic transformation

- $\mathfrak{M}(\mathrm{CIR})$ – confluent hypergeometric family (class of driftless processes related to CIR process by a stochastic transformation)

- $\mathfrak{M}(\mathrm{Jacobi})$ – hypergeometric family (class of driftless processes related to Jacobi process by a stochastic transformation)

- X_t^δ, $X_t^{h,\delta}$ and Y_t^δ – lattice approximations to X_t, X_t^h and Y_t

- $G[r, t, \lambda]$ and $F[r, t, \lambda]$ - generating and characteristic functions of affine process

- $W_{f_1, f_2}(x)$ - Wronskian of two solutions f_1, f_2 of the second order linear ODE $af'' + bf' + cf = 0$.

- $\{z, x\}$ – Schwarzian derivative of z with respect to x

- $P_2(z) = \{a_2 z^2 + a_1 z + a_0\}$ – all polynomials in z of degree 2 or less.

Introduction

The success of pricing models is often measured by the extent to which closed form solutions of the Black-Scholes type are available for the basic payoffs. Analytic tractability is often crucial for the calibration to market data and is helpful for the implementation of numerical algorithms for exotics. Extensions of the Black-Scholes formula have been directed towards three main model classes: (i) state dependent volatility models postulate a deterministic relationship between the underlying state variable, time and the local volatility; (ii) stochastic volatility models assume that the volatility follows a distinct but correlated process; (iii) and finally jump models, with state dependent or even stochastic measure of jumps.

As Dupire [6] demonstrated, state dependent volatility models are able to reproduce arbitrage-free implied volatility surfaces. Robust estimations however require either regularizations or settling on a parametric form for the local volatility. Parametric extensions of the Black-Scholes model based on geometric Brownian motion led at first to two three-parameter families, namely the CEV models (constant elasticity of variance) by Cox and Ross [5] and the quadratic volatility models in Rady [21]. On the jump-models front, geometric Brownian motions and the corresponding pricing formulas were extended in various ways; particularly noteworthy is the variance-gamma model by Madan et al. in [18], [17], [16], a class of pure jump processes admitting the Black-Scholes case as a limit, which is analytically solvable in the case of European calls and puts. Stochastic volatility models were considered by several authors including Hull-White [12] and Heston [10], [11] and Stein-Stein [7].

These three volatility models are all able to reproduce most of the features of market implied volatility skews for European options at a fixed maturity, although sometimes they require unnatural choices of parameters. State dependent volatility models run into dynamic inconsistencies, stochastic volatility models encounter difficulties capturing the very short term behavior of the implied volatility skew, which imply very large values of the instantaneous volatility as a way to emulate a jump component. Jump models have an opposite shortcoming as the corresponding implied volatility surface flattens out too rapidly unless jump amplitudes increase as a function of time at a fairly steep rate.

Studies by Bates [3] indicate that stochastic volatility and jumps are both features of the real-world process and both effects are reflected in option prices. An important feature of equity price processes is also the so-called *leverage effect,* according to which price levels are negatively correlated with spot volatility. From the modelling perspective, this effect can be captured by both local volatility models and stochastic volatility models in which equity returns are negatively correlated to volatility changes.

Motivated by these reasons, in this thesis we introduce a novel approach aimed at combining the three volatility models together within an analytically amenable framework.

Outline of the method

Let $X_t \in D_x$ be a stationary Markov process under the measure P with the generator $\mathcal{L}_X$. We say that the process X_t is *solvable* if its transitional probability density $p_X(t, x_0, x_1)$ can be computed explicitly (for example as an expansion in orthogonal polynomials, see section 2.1). Function $p_X(t, x_0, x_1)$ is a solution to the *backward Kolmogorov equation*

$$\frac{\partial p}{\partial t} + \mathcal{L}_X p = 0, \;\; t \geq 0,$$

with the initial condition $p(0, x_0, x_1) = \delta(x_1 - x_0)$ (and some boundary conditions, defined by the boundary behavior of the process X_t). Thus we could equivalently define a solvable process as a Markov process for which the backward Kolmogorov equation is solvable.

Our method of constructing solvable processes consists of two main steps: First, we start with a solvable process X_t and with the help of *stochastic transformations* we transform the process X_t into a solvable driftless process Y_t. In the second step we introduce jumps and stochastic volatility (while preserving the solvability of the model) by using a particular type of *stochastic time change* $\tilde{Y}_t = Y_{T_t}$. This combined method gives parametric families of driftless processes with up to 7 parameters for the process Y_t plus several free parameters for the choice of the time change process T_t.

Remark. The key point is that the stochastic transformation can be applied to *any* solvable process X_t, while the second step requires further restrictions on the process X_t – as we will see later we will need to be able to express the probability density of X_t as an expansion in some kind of orthogonal basis to be able to introduce time change and preserve solvability.

Stochastic transformations

First we will describe our results concerning stochastic transformations. The idea (which was first introduced in [2]) is the following: we start with a solvable process X_t and transform it into another solvable (Markov, stationary) process Y_t. To make the problem well defined we require the transformed process Y_t to be driftless. This restriction does not actually affect the generality of the method, as the drift can be introduced later (preserving solvability) by either measure change or diffeomorphism.

As we discuss in the introduction to chapter 3, the most general transformations which preserve the solvability are:

change of measure (given by Doob's h-transform, see section (3.2.1)):

$$(X_t, \mathrm{P}) \mapsto (X_t, \mathrm{P}^h), \text{ such that } \frac{d\mathrm{P}^h_t}{d\mathrm{P}_t} = e^{-\rho t} h(X_t), \tag{0.0.1}$$

where h is the ρ-excessive function (which will be constructed later)

and change of phase space:

$$X_t \mapsto Y_t = Y(X_t), \tag{0.0.2}$$

where $Y(x): \ D_x \mapsto D_y$ is a diffeomorphism (or, more generally, some invertible measurable function). (As we mentioned above, the stochastic time change will be discussed later.)

We define the *stochastic transformation* to be a triple $\{\rho, h, Y\}$, where $\rho \in \mathbb{R}$, $h = h(x)$ is a ρ-excessive function defining the measure change and $Y = Y(x)$ is a diffeomorphism, such that the process $Y(X_t)$ is driftless under the new measure P^h.

Our main result is the theorem (3.2.1), which gives an explicit method for constructing all stochastic transformations in the case when the process X_t is a diffusion process:

Theorem:

Let X_t be a stationary Markov diffusion process under the measure P on the domain $D_x \subseteq \mathbb{R}$ and admitting a Markov generator $\mathcal{L}_X$ of the form

$$\mathcal{L}_X f(x) = b(x)\frac{df(x)}{dx} + \frac{1}{2}\sigma^2(x)\frac{d^2 f(x)}{dx^2}. \tag{0.0.3}$$

We assume that both boundaries of D_x are inaccessible and that ρ is nonnegative.

Then $\{\rho, h, Y\}$ *is a stochastic transformation if and only if*

$$\begin{cases} h(x) = c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x), \\ Y(x) = \frac{1}{h(x)}(c_3\varphi_\rho^+(x) + c_4\varphi_\rho^-(x)), \end{cases}$$

where $c_i \in \mathbb{R}$ *are parameters ,*$c_1, c_2 \geq 0$*,* $c_1c_4 - c_2c_3 \neq 0$*, and* $\varphi_\rho^+(x)$ *and* $\varphi_\rho^-(x)$ *are increasing and decreasing solutions to the differential equation*

$$\mathcal{L}_X\varphi(x) = \rho\varphi(x).$$

Though we give a rigorous prove of this theorem only in the diffusion case (and in the similar case of birth and death processes), the method of constructing stochastic transformations described in this theorem can be generalized to be applicable to a bigger class of Markov processes, including processes with jumps and multidimensional processes. The method (see section 3.2.3) consists of the following three main steps:

(i) Fix $\rho \geq 0$ and find two linearly independent solutions φ_1 and φ_2 to the "eigenvalue equation"

$$\mathcal{L}_X\varphi = \rho\varphi. \tag{0.0.4}$$

For positive ρ and in the case X_t is a diffusion or birth and death process these solutions are related to the Laplace transform of the first hitting times, see lemmas (1.1.6) and (1.2.1). In these cases equation (0.0.4) is either an ODE or a finite difference equation.

(ii) Take two positive constants c_1 and c_2 and define the function

$$h(x) := c_1\varphi_1(x) + c_2\varphi_2(x).$$

The process $e^{-\rho t}h(X_t)$ is a positive local martingale, thus it is a supermartingale and we can use it to define an absolutely continuous measure change, given by equation (0.0.1).

(iii) Define function $Y(x)$:

$$Y(x) := \frac{c_3\varphi_1(x) + c_4\varphi_2(x)}{h(x)},$$

where c_3 and c_4 are some constants such that $c_1c_4 - c_2c_3 \neq 0$. It is quite easy to prove that process $Y(X_t)$ is driftless under the new measure P^h, the proof is quite general and does not use the particular form of the Markov generator $\mathcal{L}_X$ (see section 3.2.3). However it is more complicated to prove that the function $Y(x)$ is invertible – we were able to prove this only in two cases: when X_t is a diffusion and under some additional technical conditions in the case when X_t is a birth and death process, in both cases the particular form of $\mathcal{L}_X$ was important.

In what follows we study in detail the properties of transformed process Y_t in the case when X_t is a diffusion process. In lemmas (3.2.6) and (3.3.1) we compute the main characteristics of Y_t, such as *speed measure, killing measure, Green function* and most important, the probability density $p_Y(t, y_0, y_1)$. We prove that the transformed process Y_t is a transient driftless process, its domain is either infinite interval of the form $[D^1, \infty)$ or $(-\infty, D^2]$ if one of c_1, c_2 is zero, or it is a bounded interval if both constants c_1 and c_2 are not zero (see lemma (3.3.2)). In the second case the process Y_t is either a bounded (thus uniformly integrable) martingale or it is a nonconservative process (as we will see later both situations arise in applications).

Next we study the properties of stochastic transformations: in lemmas (3.3.5) and (3.3.7) we prove that stochastic transformations naturally define an equivalence relation on the set of all one-dimensional diffusions:

$$X_t \sim Y_t \quad \text{if } Y_t \text{ can be obtained from } X_t \text{ by some stochastic transformation.}$$

Thus for every process X_t we can define an equivalence class $\mathfrak{M}(X)$ – class of processes , which are related to X_t by some stochastic transformations. In lemma (3.3.8) we give an answer to the following natural question: given a diffusion process X_t with generator $\mathcal{L}_X$ and a driftless diffusion process Y_t defined by volatility function $\sigma_Y(y)$, what are conditions on $\sigma_Y(y)$ and $\mathcal{L}_X$, such that Y_t is related to X_t by a stochastic transformation?

All the results described above were quite general and could be applied to arbitrary diffusion processes. The important feature of our method is that it relies on an initial solvable Markov process X_t which we use as the starting point (as we will see later, such important properties of processes in $\mathfrak{M}(X)$ as boundary behavior depend on the initial process X). Motivated by this reasoning, in chapter 2 we give examples of solvable processes which can be used as a starting point for stochastic transformations. These processes have an important property that the *generator of X_t admits a family of orthogonal polynomials as a complete set of its eigenfunctions*, or in other words the transition probability density can be expressed as the series in orthogonal polynomials:

$$p(t,x_0,x_1)=\sum_{n=0}^{\infty}\frac{e^{\lambda_n t}}{(\psi_n,\psi_n)_m}\psi_n(x_0)\psi_n(x_1), \tag{0.0.5}$$

see section (2.1) for details. This fact will be crucial when we introduce the concept of time change later in chapter 5: the fact that the time and space variables are now conveniently separated in formula (0.0.5) will enable us to compute explicitly the transition probability density for the time changed process.

We use these processes (Ornstein-Uhlenbeck (OU), CIR and Jacobi), to model the process X_t, and then applying stochastic transformations we construct the class $\mathfrak{M}(X)$ of driftless processes. Our next results are concerned with the properties of the processes in $\mathfrak{M}(X)$.

In lemma (3.4.6) we prove that processes in $\mathfrak{M}(\text{OU})$ are conservative with both boundaries being natural (thus they are bounded martingales if $c_1>0$ and $c_2>0$). In lemmas (3.4.7) and (3.4.10) we show that processes in $\mathfrak{M}(\text{CIR})$ have one natural boundary and one killing or exit boundary depending on the parameters, and processes in $\mathfrak{M}(\text{Jacobi})$ have both boundaries killing or exit. In each case we find explicit expressions for functions φ^+_λ, φ^-_λ, which in turn enables us to compute functions h and $Y(x)$ (and thus $X(y)=Y^{-1}(y)$) and to use these processes in applications.

We see that by using one of these processes we can cover any boundary behavior – one can have a pure martingale (associated to OU process), a process with one killing/exit and one natural boundary (associated to CIR process) or a process with both boundaries being killing/exit (associated to Jacobi process). In some sense these are the most general families: it was proved in [20], that OU,CIR and Jacobi are the only diffusion processes associated with a family of orthogonal polynomials.

By using birth and death analogs of above processes (Charlier, Meixner and Hahn) one can construct driftless processes on the lattice (see section 3.5), which can be considered lattice approximations to the corresponding diffusion processes, and thus are very important for applications (for example for pricing American style options, which will be considered later).

Classification of solvable diffusions

Our next results concern the problem of classification of solvable diffusions (see chapter 4). We consider the class of diffusion processes, for which the transition probability density can be computed as an integral over hypergeometric functions. As we show it is equivalent to the following condition: all the solutions to the “eigenvalue equation” (0.0.4) can be expressed in terms of hypergeometric function. This class of processes seems to be the most general, as it includes all the processes constructed above. Using tools from the theory of ODEs, such as Liouville transformations, canonical forms of second order ODEs and Bose invariants, we were able to prove the following two classification results:

First classification theorem:

A driftless process Y_t is solvable in the sense of definition (4.1.1) if and only if its volatility function is of the following form:

$$\sigma_Y(y) = \sigma_Y(Y(x)) = C\sqrt{A(x)}\frac{W(x)}{(c_1F_1(x)+c_2F_2(x))^2}\sqrt{\frac{A(x)}{R(x)}}, \tag{0.0.6}$$

where change of variables is given by

$$y = Y(x) = \frac{c_3F_1(x)+c_4F_2(x)}{c_1F_1(x)+c_2F_2(x)}, \quad c_1c_4 - c_3c_2 \neq 0. \tag{0.0.7}$$

In the case $A(x) = x$:

(i) $R(x) \in P_2$, such that $R(x) \neq 0$ in $(0, \infty)$

(ii) F_1 and F_2 are functions $M(a, b, wx)$ and $U(a, b, wx)$

(iii) $W(x)$ is a Wronskian of the scaled Kummer differential equation (equal to $s'(x)$ in (2.2.16)).

and in the case $A(x) = x(1-x)$

(i) $R(x) \in P_2$, such that $R(x) \neq 0$ in $(0, 1)$

(ii) F_1 and F_2 are two linearly independent solutions to the hypergeometric equation given by (1.3.11).

(iii) $W(x)$ is a Wronskian of the hypergeometric differential equation (equal to $s'(x)$ in (2.2.29)).

Second classification theorem:

Let $R(x) \in P_2$ be a second degree polynomial in x.

(i) In the confluent hypergeometric case: Assume that $R(x)$ has no zeros in $(0, \infty)$. Let $X_t = X_t^R$ be the diffusion process with dynamics:

$$dX_t = (a + bX_t)\frac{X_t}{R(X_t)}dt + \frac{X_t}{\sqrt{R(X_t)}}dW_t \tag{0.0.8}$$

Then confluent hypergeometric R-family coincides with $\mathfrak{M}(X_t^R)$ and thus can be obtained from X_t^R by stochastic transformations. In the particular case $R(x) = A(x) = x$ we obtain the CIR family defined in (3.4.4).

(ii) *In the hypergeometric case: Assume that $R(x)$ has no zeros in $(0,1)$. Let $X_t = X_t^R$ be the diffusion process with dynamics:*

$$dX_t = (a + bX_t)\frac{X_t(1-X_t)}{R(X_t)}dt + \frac{X_t(1-X_t)}{\sqrt{R(X_t)}}dW_t \tag{0.0.9}$$

Then hypergeometric R-family coincides with $\mathfrak{M}(X_t^R)$ and thus can be obtained from X_t^R by stochastic transformations. In the particular case $R(x) = A(x) = x(1-x)$ we obtain the Jacobi family defined in (3.4.5).

Stochastic time change

Our second major step in constructing solvable Markov processes is achieved by introducing the stochastic time change:

$$t \mapsto T_t, \;\; Y_t \mapsto \tilde{Y}_t = Y_{T_t}, \tag{0.0.10}$$

where T_t is a time change process – increasing right-continuous process with $T_0 = 0$. *We make an important assumption that the process T_t is independent of Y_t.* The main result is lemma (5.2.3):

Lemma: *Let the process X_t be associated with a family of orthogonal polynomials $\psi_n(x)$ and assume that Y_t is obtained by a stochastic transformation $\{\rho, h, Y\}$ from the process X_t. Then the probability density for the process $\tilde{Y}_t = Y_{T_t}$ can be computed as*

$$p_{\tilde{Y}}(t, y_0, y_1) = \frac{1}{h(x_0)h(x_1)}\sum_{n=0}^{\infty} L(\rho - \lambda_n, t)\psi_n(x_0)\psi_n(x_1), \tag{0.0.11}$$

where $y_i = Y(x_i)$ and $L(\lambda, t) = Ee^{-\lambda T_t}$ is the Laplace transform of the random variable T_t.

As the above lemma shows, the process $\tilde{Y}_t$ is solvable if and only if one can compute the Laplace transform of T_t explicitly. This restriction still leaves us with quite a big class of stochastic time change processes. We propose to model the time change process as

$$T_t = f(t) + T_t^c + \gamma_t,$$

where $f(t)$ is an increasing right-continuous deterministic function, T_t^c is an absolutely continuous process independent of Y_t ($T_t^c = \int_0^t r_s ds$ for some positive process r_s) and γ_t is a pure jump process, modelled by an increasing Levy process (independent of r_t and Y_t). Since all the three components are independent, the Laplace transform for this type of time change process can be computed as the product of the corresponding Laplace transforms of each component (see section 5.2.3).

To model the process γ_t we choose either the *Gamma process* or the *stable process* due to the fact that the Laplace transform of γ_t can be computed explicitly in these cases (see section 5.2.1). To model the process r_s we use the so-called *affine processes,* for which the Laplace transform of T_t^c can be computed as

$$L_{T^c}(\lambda, t) = E(e^{-\lambda \int_0^t r_s ds} | r_0) = e^{m(t,\lambda) r_0 + n(t,\lambda)},$$

for some deterministic functions m and n. Two well known examples are OU and CIR processes (see [15]). In section 5.1 we prove that the birth and death approximations to OU and CIR given by Charlier and Meixner processes are also affine and we find the explicit expressions for functions m and n. The Meixner process will be extremely important when we will use it to model the stochastic volatility of the driftless process in our later application to American style options. In the section 5.1.1 we also present an algorithm for computing the discounted transitional probabilities, defined as

$$q_r(t, j, k) = E\left(\mathbb{I}(r_t = k\delta) e^{-\int_0^t r_s ds} | r_0 = j\delta\right).$$

It is easy to understand the effect of the stochastic time change $Y_t \mapsto \tilde{Y}_t = Y_{T_t}$, where $T_t = f(t) + T_t^c + \gamma_t$: the components $f(t)$ and T_t^c bring state dependent and stochastic volatility into our model, while γ_t is responsible for jumps. Thus we have constructed (multi parameter) families of solvable driftless stationary Markov processes with state dependent volatility, stochastic volatility and jumps.

Applications of solvable processes

Our next results are concerned with possible applications of the solvable processes constructed above to the Mathematical Finance. As the first application we use the process $\tilde{Y}_t$ to model the stock price in the spot market. The process $\tilde{Y}_t$ is a solvable driftless process, which is constructed as a time changed process $\tilde{Y}_t = Y_{T_t}$, where Y_t is related by a stochastic transformation $\{\rho, h, Y\}$ to one of the processes described in chapter 2 (processes which are associated with a family of orthogonal polynomials $\psi_n(x)$).

For simplicity we assume that the interest rate is zero (the banking account $B_t \equiv 1$). One can easily generalize the model to include constant interest rate $r_t \equiv r$ for pricing European style calls and puts.

We prove the following lemma:

Lemma: *The price of the call option with an underlying $\tilde{Y}$ can be computed as:*

$$C_{\tilde{Y}}(t, y_0, K) := E^{\mathrm{P}^h}((\tilde{Y}_t - K)^+ | Y_0 = y_0)$$
$$= (y_0 - K)^+ + \frac{1}{2}\sigma_Y^2(K)\frac{m_Y(K)}{h(x_0)h(k)}\left[G_X(\rho, x_0, k) + \sum_{n=0}^{\infty}\frac{L(t, \rho - \lambda_n)}{\rho - \lambda_n}\psi_n(x_0)\psi_n(k)\right],$$

where $K = Y(k)$ and $y_0 = Y(x_0)$.

In this formula G_X is the Green function of the process X_t (see chapter 1), $m_Y(y)$ and $\sigma_Y(y)$ are the speed measure and volatility of the process Y_t (see lemma (3.3.1)), $L(\lambda, t)$ is the Laplace transform of the time change process T_t. It is interesting to note that a similar formula can be proved for the case of birth and death processes (see lemma (5.3.5)). This formula is useful since it allows us to calibrate our model efficiently to the real market data.

As another possible application of solvable processes to Mathematical Finance, we illustrate an example of pricing an American style options with stochastic volatility and jumps. Again we model the stock price as a solvable process with stochastic volatility and jumps $\tilde{Y}_t$ and assume that the banking account $B_t \equiv 1$. We use the Hahn process (see section (2.2.6)) to model X_t, then we use the stochastic transformation to construct Y_t and finally the time change $t \mapsto T_t = f(t) + T_t^c + \gamma_t$ to arrive at the process $\tilde{Y}_t$. We choose to use the Hahn process due to the obvious advantage that there is only a finite number of Hahn polynomials (the Hahn process is a finite Markov chain), thus all the expansions in orthogonal polynomials are just finite sums.

To price American style options one could use the well known backward recursion algorithm (see section 5.4 or [23]). However difficulties arise here since the process $\tilde{Y}_t$ is clearly not Markov (since it has "stochastic volatility" r_t), thus we have to consider instead the two dimensional process $(r_t, \tilde{Y}_t)$ (which is now Markov). The following formula for computing the two-dimensional transition probabilities for the process $(r_t, \tilde{Y}_t)$ is our main result in this section:

Lemma: *The transition probabilities of the Markov process* $(r_t, \tilde{Y}_t)$ *can be computed as:*

$$p_{\tilde{Y},r}(t, y_0, y_1; r_0, r_1) = m(x_1)\frac{h(x_1)}{h(x_0)}\sum_{n=0}^{\infty} l_n(t, r_0, r_1)\psi_n(x_0)\psi_n(x_1) \quad (0.0.12)$$

where the factor $l_n(t, r_0, r_1)$ *is given by*

$$l_n(t, r_0, r_1)) = e^{-(\rho-\lambda_n)f(t)} L_\gamma(\rho - \lambda_n, t) q_r(\rho - \lambda_n, t, r_0, r_1), \quad (0.0.13)$$

and $q_r(\rho - \lambda_n, t, r_0, r_1)$ *are the discounted transition probabilities of the process* $(\rho - \lambda_n)r_s$*:*

$$q_r(\rho - \lambda_n, t, r_0, r_1) = E\left(e^{-(\rho-\lambda_n)\int_0^t r_s ds} | r_t = r_1, r_0\right). \quad (0.0.14)$$

Outline of the thesis

The thesis is organized as follows:

In the first chapter we present necessary definitions, facts and theorems on which we will rely in later chapters. The first section is dedicated to the classical theory of one-dimensional diffusions: construction and probabilistic description of such important analytical tools as speed measure, scale function, Green function, etc., description of the boundary behavior of diffusion process X_t. In the second section we briefly introduce birth and death processes and in the last section we present the necessary definitions and facts about hypergeometric and confluent hypergeometric functions, which will be used later in chapters 3 and 4.

In the second chapter we present the necessary facts about OU (Ornstein-Uhlenbeck), CIR (Cox-Ingersoll-Ross) and Jacobi processes and their lattice approximations given by birth and death processes (Charlier, Meixner and Hahn processes). These particular processes are chosen because they are associated with families of orthogonal polynomials and will be used as a starting point for constructing solvable driftless processes in the next chapters. The main objective in this chapter is to express in each case the transition probability density $p(t, x_0, x_1)$ as an expansion in orthogonal polynomials and to study the boundary behavior of these processes.

Chapter 3 is dedicated to stochastic transformations, as a way to obtain new (stationary) driftless solvable processes from already known ones. The main result is theorem (3.2.1), which gives an explicit algorithm to construct all stochastic transformations. In what follows we study the properties of transformed processes, show that such well known examples as the quadratic volatility family and CEV processes can be easily obtained by stochastic transformations. We construct three new families of solvable processes related to OU,CIR and Jacobi processes and generalize this method to birth and death processes, which can be considered as lattice approximations to the processes in continuous phase space.

In chapter 4 we classify all one-dimensional driftless Markov process for which the transition probability density can be computed as an integral over hypergeometric functions. Theorem (4.3.1) finds an explicit form of the volatility functions of these processes while theorem (4.3.7) gives a different construction of these processes using stochastic transformations. This is an analog of construction in the previous chapter, where we used OU, CIR and Jacobi processes as starting points to construct new families of driftless processes.

In the last chapter we apply the theory developed above to several problems in Mathematical Finance. In the first section we prove that Meixner and Charlier processes (which are lattice approximations to CIR and OU) are also affine, and we give a method of computing the (discounted) transitional probabilities as Fourier transform of generating function (for which we find an explicit expression). Then we discuss the concept of time change, and we show how one can introduce jumps and stochastic volatility in the state-dependent volatility models constructed in chapter 3. In the next section we use the driftless process with stochastic volatility and jumps to model the stock price under the risk neutral measure and we find an explicit expression for the price of the European call option in this model (also in the case of process on the lattice). In the last section we present an algorithm to price American style options with stochastic volatility and jumps. We model the stock price as a driftless process on the lattice, with jumps and stochastic volatility, where the volatility process is modelled by a Meixner process. We use the recursive procedure with a two-dimensional lattice to find the price of an American option and its optimal exercise boundary.

Chapter 1

Background

This is an introductory chapter, in which we give necessary definitions, facts and theorems which are used in later chapters. The first section is dedicated to the classical theory of one-dimensional diffusion: construction and probabilistic descriptions of such important analytical tools as speed measure, scale function, Green function, etc., and the description of the boundary behavior of the diffusion process X_t. For good references on this subject see [19], [13], [24] and [4].

In the second section we briefly introduce birth and death processes (see [25]) and we give one result about the Laplace transform of the first exit time, which we need later in chapter 3.

In the last section we present definitions and facts about hypergeometric and confluent hypergeometric functions, which are used in chapters 3 and 4.

1.1 Diffusion processes

In this section we will not use measure change or any other transformations of stochastic processes, thus we will simplify notations and will not write subscripts specifying the process and/or the measure.

Let D be the interval $[D^1, D^2] \subseteq \mathbb{R}$. Let X_t be the *stationary Markov process* taking values in D with transition function $P(t, x, A) = \mathrm{P}_x(X_t \in A)$. The *probability semigroup* is defined as

$$P(t)f(x) = \int f(y)P(t, x, dy) \tag{1.1.1}$$

and the *resolvent operator* is

$$R(\lambda)f(x) = \int_0^\infty e^{-\lambda t} P(t)f(x)dt \tag{1.1.2}$$

for bounded measurable functions $f : D \to \mathbb{R}$. The process X_t is called *conservative* if $P(t,x,D) \equiv 1$ for all t and all $x \in D$. If the process is not conservative we can enlarge the state space by adding a *cemetery point* Δ_∞:

$$P(t,x,\Delta_\infty) = 1 - P(t,x,D).$$

Then it is easy to see that the process X_t on the space $D \cup \Delta_\infty$ is conservative. If f is a function on D, we will extend it to $D \cup \Delta_\infty$ by letting $f(\Delta_\infty) = 0$.

The *(strong) infinitesimal generator* $\mathcal{L}$ of the process X_t is defined as:

$$\mathcal{L}f := \frac{d}{dt}P(0)f = \lim_{t\to 0}\frac{P(t)f - f}{t} \tag{1.1.3}$$

for all continuous, bounded $f : D \to \mathbb{R}$, such that the limit exists in the norm. The set of all these f is called *the domain of* $\mathcal{L}$ and is denoted $\mathcal{D}(\mathcal{L})$.

From now on we will assume that X_t is a regular diffusion process, specified by its Markov generator

$$\mathcal{L}f = \frac{1}{2}\sigma^2(x)f''(x) + b(x)f'(x) - c(x)f(x) \tag{1.1.4}$$

where functions $b(x)$,$c(x)$ and $\sigma(x)$ are smooth and $c(x) \geq 0$, $\sigma(x) > 0$ in the interior of D.

Every diffusion process has three basic characteristics: *speed measure* $m(dx)$, *scale function* $s(x)$ and *killing measure* $k(dx)$. For the diffusion specified by the generator (1.1.4) speed measure and killing measure are absolutely continuous with respect to Lebesgue measure (in the interior of domain D):

$$m(dx) = m(x)dx, \quad k(dx) = k(x)dx,$$

and functions $m(x)$, $k(x)$ and $s(x)$ are defined as follows:

$$m(x) = 2\sigma^{-2}(x)e^{B(x)}, \quad s'(x) = e^{-B(x)}, \quad k(x) = c(x)m(x) = 2c(x)\sigma^{-2}(x)e^{B(x)} \tag{1.1.5}$$

where $B(x) := \int^x 2\sigma^{-2}(y)b(y)dy$.

Remark 1.1.1. We denote by $m(x)$ a density $\frac{m(dx)}{dx}$. The same applies for the killing measure $k(dx)$.

The probabilistic interpretation of m, s and k is the following:

- Assume $k \equiv 0$. Let $H_z := \inf\{t : X_t = z\}$ and $(a,b) \subset D$. Then

  $$\mathrm{P}_x(H_a < H_b) = \frac{s(b) - s(x)}{s(b) - s(a)}.$$

  We say that X_t is in *natural scale* if $s(x) = x$. In this case (if the process is conservative) X_t is a local martingale.

- The speed measure is characterized by the following: for every $t > 0$ and $x \in D$, the transition function $P(t, x, dy)$ is absolutely continuous with respect to $m(dy)$:

$$P(t, x, A) = \int_A p(t, x, y) m(dy)$$

and the density $p(t, x, y)$ is positive, jointly continuous in all variables and symmetric: $p(t, x, y) = p(t, y, x)$.

In the case X_t is in natural scale, the following is true: if the speed measure $m(x)$ is ”large”, the increment $X(t + h) - X(t)$ is ”small”.

- The killing measure is associated to the distribution of the location of the process at its lifetime $\zeta := \inf\{t : X_t \notin D\}$:

$$\mathrm{P}_x(X_{\zeta-} \in A | \zeta < t) = \int_0^t ds \int_A p(s, x, y) k(dy).$$

From now on we will assume that there is no killing in the interior of domain D, that is function $c(x) \equiv 0$.

Remark 1.1.2. Note that the scale function can be characterized as a solution to equation

$$\mathcal{L} s(x) = 0,$$

and $s'(x)$ is proportional to the Wronskian $W_{\varphi_1, \varphi_2}(x)$, where φ_1, φ_2 are any two linearly independent solutions to

$$\mathcal{L} \varphi = \lambda \varphi.$$

Let τ be the stopping time with respect to filtration $\{\mathcal{F}_t\}$. The process $X_{\tau \wedge t}$ is called *the process stopped at* τ and is denoted by X_t^τ.

We will need the following lemma in later chapters:

Lemma 1.1.3. *Let* $T = \inf\{t \geq 0 : \ X_t \notin \mathrm{int}(D)\}$ *be the first time the process* X_t *hits the boundary of* D. *Then for each* $x \in D$, $Y_t^T \equiv s(X_t^T)$ *is a continuous* P_x*-local martingale.*

For the proof of this lemma see [24], vol. II, p.276.

The generator $\mathcal{L}$ of the process X_t can be expressed as

$$\mathcal{L}f = \frac{d}{m(dx)}\frac{df(x)}{ds(x)} = D_m D_s f, \tag{1.1.6}$$

thus the speed measure and the scale function determine the process X_t up to the first time it hits the boundary of the interval.

Remark 1.1.4. Note that equation 1.1.6 can serve as a more general definition of the speed measure and scale function (see [4]) and can be applied to the case when the speed measure is not absolutely continuous with respect to Lebesgue measure.

The boundary behavior of the process X_t is described by the following classical result (see [13],[19]):

Lemma 1.1.5. Feller classification of boundary points. *Let $d \in (D^1, D^2)$. Define functions $R(x) = m((d,x))s'(x)$ and $Q(x) = s(x)m(x)$. Fix small $\epsilon > 0$ (such that $D^1 + \epsilon \in D$). Then the endpoint D^1 is said to be:*

$$\begin{cases} \textit{regular if} & Q \in L^1(D^1, D^1+\epsilon), \quad R \in L^1(D^1, D^1+\epsilon) \\ \textit{exit if} & Q \notin L^1(D^1, D^1+\epsilon), \quad R \in L^1(D^1, D^1+\epsilon) \\ \textit{entrance if} & Q \in L^1(D^1, D^1+\epsilon), \quad R \notin L^1(D^1, D^1+\epsilon) \\ \textit{natural if} & Q \notin L^1(D^1, D^1+\epsilon), \quad R \notin L^1(D^1, D^1+\epsilon) \end{cases} \tag{1.1.7}$$

The same holds true for D^2.

Regular or exit boundaries are called *accessible*, while entrance and natural boundaries are called *inaccessible*.

An *exit* boundary can be reached from any interior point of D with positive probability. However it is not possible to start the process from an exit boundary.

The process cannot reach an *entrance* boundary from any interior point of D, but it is possible to start the process at an entrance boundary.

A *natural* boundary cannot be reached in finite time and it is impossible to start a process from the natural boundary. The natural boundary D_1 is called *attractive* if $X_t \to D^1$ as $t \to \infty$.

A *regular* boundary is also called *non-singular*. A diffusion reaches a non-singular boundary with positive probability. In this case the characteristics of the process do not determine the process uniquely and one has to specify boundary conditions at each non-singular boundary point: if $m(\{D^i\}) < \infty$, $k(\{D^i\}) < \infty$, then we have the boundary conditions:

$$\begin{cases} g(D^1)m(\{D^1\}) - \frac{df(D^1)}{ds(x)} + f(D^1)k(\{D^1\}) = 0, \\ g(D^2)m(\{D^2\}) + \frac{df(D^2)}{ds(x)} + f(D^2)k(\{D^2\}) = 0. \end{cases} \tag{1.1.8}$$

where $g := \mathcal{L}f$ for $f \in \mathcal{D}(\mathcal{L})$.

The following terminology is used: the left endpoint D^1 is called

- *reflecting*, if $m(\{D^1\}) = k(\{D^1\}) = 0$,
- *sticky*, if $m(\{D^1\}) > 0$, $k(\{D^1\}) = 0$,
- *elastic*, if $m(\{D^1\}) = 0$, $k(\{D^1\}) > 0$.

A diffusion process spends no time and does not die at a reflecting boundary point. X does not die, but spends a positive amount of time at a sticky point (which in the case $m(\{D^1\}) = \infty$ is called an *absorbing boundary* - the process stays at D^1 forever after hitting it). X does not spend any time at elastic boundary - it is either reflected or dies with positive probability after hitting D^1 (in the limit $k(\{D^1\}) = \infty$ we call D^1 a *killing boundary*, since that X is killed immediately if it hits D^1).

Let the interval D be an infinite interval, for example of the form $[D^1, \infty)$. We say that the process X_t *explodes* if the boundary $D^2 = \infty$ is an accessible boundary. Using the previous lemma one can see that the process explodes if and only if for some $\epsilon > 0$

$$R(x) = m((D^1 + \epsilon, x))s'(x) \in L^1(D^1 + \epsilon, \infty). \tag{1.1.9}$$

In chapter 3 we need to construct two linearly independent solutions to the ODE

$$\mathcal{L}\varphi(x) = \lambda\varphi(x), \quad \lambda > 0, \quad x \in D. \tag{1.1.10}$$

The probabilistic description of these solutions is given by the following lemma (see [24], vol. II, p. 292):

Lemma 1.1.6. *For $\lambda > 0$ there exist an increasing $\varphi_\lambda^+(x)$ and a decreasing $\varphi_\lambda^-(x)$ solutions to equation (1.1.10). These solutions are convex, finite in the interior of the domain D and are related to the Laplace transform of the first hitting time H_z as follows:*

$$E_x\left(e^{-\lambda H_z}\right) = \begin{cases} \frac{\varphi_\lambda^+(x)}{\varphi_\lambda^+(z)}, & x \le z, \\ \frac{\varphi_\lambda^-(x)}{\varphi_\lambda^-(z)}, & x \ge z. \end{cases} \tag{1.1.11}$$

The following theorem due to W. Feller characterizes boundaries in terms of solutions to the equation (1.1.10):

Theorem 1.1.7. *(i) The boundary point D^2 is regular if and only if there exist two positive, decreasing solutions φ_1 and φ_2 of (1.1.10) satisfying*

$$\lim_{x \to D^2} \varphi_1(x) = 0, \quad \lim_{x \to D^2} \frac{d\varphi_1(x)}{ds(x)} = -1, \quad \lim_{x \to D^2} \varphi_2(x) = 1, \quad \lim_{x \to D^2} \frac{d\varphi_2(x)}{ds(x)} = 0. \tag{1.1.12}$$

(ii) The boundary point D^2 is exit if and only if every solution of (1.1.10) is bounded and every positive decreasing solution φ_1 satisfies

$$\lim_{x \to D^2} \varphi_1(x) = 0, \quad \lim_{x \to D^2} \frac{d\varphi_1(x)}{ds(x)} \leq 0. \tag{1.1.13}$$

(iii) The boundary point D^2 is entrance if and only if there exists a positive decreasing solution φ_1 of (1.1.10) satisfying

$$\lim_{x \to D^2} \varphi_1(x) = 1, \quad \lim_{x \to D^2} \frac{d\varphi_1(x)}{ds(x)} = 0, \tag{1.1.14}$$

and every solution of (1.1.10) independent of φ_1 is unbounded at D^2. In this case no nonzero solution tends to 0 as $x \to D^2$.

(iv) The boundary point D^2 is natural if and only if there exists a positive decreasing solution φ_1 of (1.1.10) satisfying

$$\lim_{x \to D^2} \varphi_1(x) = 0, \quad \lim_{x \to D^2} \frac{d\varphi_1(x)}{ds(x)} = 0, \tag{1.1.15}$$

and every solution of (1.1.10) independent of φ_1 is unbounded at D^2.

In cases (i) and (ii), all solutions of (1.1.10) are bounded near D_2 and there is a positive increasing solution z such that $\lim_{x \to D^2} z(x) = 1$. In cases (iii) and (iv) every positive, increasing solution z satisfies $\lim_{x \to D^2} z(x) = \infty$.

The following lemma is of crucial importance in chapter 3:

Lemma 1.1.8. *Let $X_0 > z$, $\rho > 0$ and $\tau = H_z$ be the first hitting time of z. Then the process $e^{-\rho t} \varphi_\rho^-(X_t)$ stopped at $t = \tau$ is a martingale. If $X_0 < z$ the same is true for function φ_ρ^+.*

Proof. From the previous lemma we find

$$\varphi_\rho^-(x) = \varphi_\rho^-(z) E_{x,0}\left(e^{-\rho\tau}\right),$$

where $E_{x,t}(\cdot) = E(\cdot|X_t = x)$. We need to show that

$$\varphi_\rho^-(x) = E_{0,x}\left(e^{-\rho t\wedge\tau}\varphi_\rho^-(X_{t\wedge\tau})\right).$$

The right hand side can be rewritten as follows:

$$E_{0,x}\left(e^{-\rho t\wedge\tau}\varphi_\rho^-(X_{t\wedge\tau})\right) = E_{0,x}\left(e^{-\rho t\wedge\tau}\varphi_\rho^-(X_{t\wedge\tau})|\tau \le t\right) + E_{0,x}\left(e^{-\rho t\wedge\tau}\varphi_\rho^-(X_{t\wedge\tau})|\tau > t\right). \quad (1.1.16)$$

Since $X_\tau = z$ a.s., the first expectation is equal $\varphi_\rho^-(z)E_{0,x}\left(e^{-\rho\tau}|\tau < t\right)$. The second expectation is equal

$$E_{0,x}\left(e^{-\rho t}\varphi_\rho^-(X_t)|\tau > t\right) = \varphi_\rho^-(z)E_{0,x}\left(e^{-\rho t}E_{0,X_t}\left(e^{-\rho\hat\tau}\right)|\tau > t\right) \quad (1.1.17)$$

where $\hat\tau$ is the first hitting time of z for the process started at time $t = 0$ at $x = X_t$. Since the process X_t is a stationary Markov process, we find that

$$E_{0,X_t}\left(e^{-\rho\hat\tau}|\tau > t\right) = E_{t,X_t}\left(e^{-\rho(\tau - t)}|\tau > t\right),$$

which together with formula (1.1.17) gives us the following equation:

$$E_{0,x}\left(e^{-\rho t}\varphi_\rho^-(X_t)|\tau > t\right) = \varphi_\rho^-(z)E_{0,x}\left(E_{t,X_t}\left(e^{-\rho\tau}\right)|\tau > t\right) = \varphi_\rho^-(z)E_{0,x}\left(e^{-\rho\tau}|\tau > t\right). \quad (1.1.18)$$

Combining formulas (1.1.18) and (1.1.16) we obtain:

$$E_{0,x}\left(e^{-\rho t\wedge\tau}\varphi_\rho^-(X_{t\wedge\tau})\right) = \varphi_\rho^-(z)E_{0,x}\left(e^{-\rho\tau}|\tau \le t\right) + \varphi_\rho^-(z)E_{0,x}\left(e^{-\rho\tau}|\tau > t\right) =$$

$$= \varphi_\rho^-(z)E_{0,x}\left(e^{-\rho\tau}\right) = \varphi_\rho^-(x),$$

which ends the proof. □

The functions $\varphi_\lambda^+(x)$ and $\varphi_\lambda^-(x)$ are also called the *fundamental solutions* of equation (1.1.10). These functions are linearly independent and their Wronskian can be computed as

$$W_{\varphi_\lambda^+,\varphi_\lambda^-}(x) = \frac{d\varphi_\lambda^+(x)}{dx}\varphi_\lambda^-(x) - \varphi_\lambda^+(x)\frac{d\varphi_\lambda^-(x)}{dx} = w_\lambda s'(x), \quad (1.1.19)$$

thus the Wronskian with respect to $D_s = d/ds(x)$ is constant:

$$W_{\varphi_\lambda^+,\varphi_\lambda^-}(x) = \frac{d\varphi_\lambda^+(x)}{ds(x)}\varphi_\lambda^-(x) - \varphi_\lambda^+(x)\frac{d\varphi_\lambda^-(x)}{ds(x)} = w_\lambda. \quad (1.1.20)$$

The *Green function* $G(\lambda, x, y)$ is defined as the Laplace transform of $p(t, x, y)$:

$$G(\lambda, x, y) := \int_0^\infty e^{-\lambda t} p(t, x, y) dt. \tag{1.1.21}$$

The Green function is symmetric and it is the kernel of the resolvent operator $R(\lambda) = (\mathcal{L} - \lambda)^{-1}$ with respect to $m(dx)$.

The Green function can be conveniently expressed in terms of functions $\varphi_\lambda^+(x)$ and $\varphi_\lambda^-(x)$ as:

$$G(\lambda, x, y) = \begin{cases} w_\lambda^{-1} \varphi_\lambda^+(x) \varphi_\lambda^-(y), & x \leq y \\ w_\lambda^{-1} \varphi_\lambda^+(y) \varphi_\lambda^-(x), & y \leq x. \end{cases} \tag{1.1.22}$$

A diffusion X_t is said to be *recurrent* if $\mathrm{P}_x(H_y < \infty) = 1$ for all $x, y \in D$. A diffusion which is not recurrent is called *transient*.

A recurrent diffusion is called *null recurrent* if $E_x(H_y) = \infty$ for all $x, y \in D$ and *positively recurrent* if $E_x(H_y) < \infty$ for all $x, y \in D$.

The following is a list of useful facts concerning recurrence, Green function and speed measure:

- X_t is recurrent if and only if $\lim\limits_{\lambda \searrow 0} G(\lambda, x, y) = \infty$.
- X_t is transient if and only if $\lim\limits_{\lambda \searrow 0} G(\lambda, x, y) < \infty$.
- X_t is positively recurrent if and only if $m(D) < \infty$. In the recurrent case we have: $\lim\limits_{\lambda \searrow 0} \lambda G(\lambda, x, y) = \frac{1}{m(D)}$ and the speed measure $m(dx)$ is a *stationary (invariant) measure* of X_t:

$$mP(t)(A) := \int_A m(dx) P(t, x, A) = m(A).$$

1.2 Birth and Death processes

Let X_t be a stationary Markov chain with values in the domain $D = \mathbb{Z}^+ = \{0, 1, 2, ..\}$ with transitional probabilities $p(t, x_0, x_1)$ and semigroup $P(t)$. The generator of the process X can be represented as a matrix $Q = (q_{ij})$: for any bounded $f \ : D \to \mathbb{R}$

$$\mathcal{L} f(i) = \sum_{k=0}^{\infty} q_{ik} f(k). \tag{1.2.1}$$

The process X is called a *birth and death process* if the matrix Q is of the following tridiagonal form:

$$Q = \begin{pmatrix} -d(0) & d(0) & 0 & \dots & 0 & 0 & 0 & \dots \\ a(1) & -(a(1)+d(1)) & d(1) & \dots & 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a(n) & -(a(n)+d(n)) & d(n) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \tag{1.2.2}$$

that is $q_{ii} = -(a(i)+d(i))$, $q_{ii+1} = d(i)$, $q_{ii-1} = a(i)$ and $q_{ij} = 0$ if $|i-j| > 1$.

One can easily check that the process X is a birth and death process if and only if its generator is a finite difference operator of the following form:

$$\mathcal{L}f(i) = (d(i)-a(i))\nabla^+ f(i) + a(i)\Delta f(i) = (d(i)-a(i))\nabla^- f(i) + d(i)\Delta f(i), \tag{1.2.3}$$

where ∇^+ (∇^-) is a forward (backward) difference operator.

A necessary and sufficient condition for the process X_t to be a birth and death process (or for the matrix Q to be of the form (1.2.2)), is that as $t \to 0$

$$p(t,i,j) = \begin{cases} d(i)t + o(t), & \text{if } j = i+1, \\ a(i)t + o(t), & \text{if } j = i-1, \\ 1-(a(i)+d(i))t + o(t), & \text{if } j = i. \end{cases}$$

The probabilistic description of X_t is the following: if at time t the process is in the state i, then independent of its past it stays there for an amount of time τ, where τ has exponential distribution with parameter $-q_{ii}$, and at time $t+\tau$ the process X_t jumps to the state $i+1$ with probability $d(i)/(-q_{ii})$ or to the state $i-1$ with probability $a(i)/(-q_{ii})$. Thus a birth and death process behaves like a diffusion – it can travel from state n to state $n+k$ only through all the states $n+1, n+2, \dots$ with probability one. We will see later that birth and death processes enjoy many of the properties of diffusions. The first example of this is the following lemma, which characterizes the solutions to equation

$$\mathcal{L}_X \varphi(x) = \lambda \varphi(x), \tag{1.2.4}$$

as Laplace transforms of the first passage time and is a discrete analog of the lemma (1.1.6):

Lemma 1.2.1. *Let $A \subset \mathbb{Z}^+$ and A is nonempty. Define*

$$\tau_A = \inf\{t \geq 0 : \ X_t \in A\} \tag{1.2.5}$$

be the first passage time to A and let

$$\varphi_{A,\lambda}(x) = E_x\left(e^{-\lambda \tau_A}\right). \tag{1.2.6}$$

Then $\varphi(x) = \varphi_{A,\lambda}(x)$ satisfies the difference equation

$$\begin{cases} \mathcal{L}_X \varphi(x) = a(x)\varphi(x-1) - (a(x)+d(x))\varphi(x) + d(x)\varphi(x+1) = \lambda\varphi(x), & \textit{if } x \notin A, \\ \varphi(x) = 1, & \textit{if } x \in A. \end{cases} \tag{1.2.7}$$

For the proof of this lemma see [25].

1.3 Hypergeometric functions

In this section, we review basic notions about hypergeometric functions and Fuchsian differential equations which will be important in the following chapters. A good collection of facts and formulas can be found in [8] and [1].

Hypergeometric functions are defined through Taylor series expansion which generalize the geometric series

$$(1-z)^{-\alpha} = \sum_{n=0}^{\infty} \frac{\alpha(\alpha+1)...(\alpha+n-1)}{n!} z^n = \sum_{n=0}^{\infty} \frac{(\alpha)_n}{n!} z^n. \tag{1.3.1}$$

Generalizing this expansion, we introduce the functions

$${}_pF_q(\alpha_1, ..., \alpha_p; \beta_1, ...\beta_q; z), \quad p \leq q+1, \beta_j \in \mathbb{C} \setminus -\mathbb{Z}_+ \tag{1.3.2}$$

which, for $|z| < 1$, is the sum of the series

$${}_pF_q(\alpha_1, ..., \alpha_p; \beta_1, ...\beta_q; z) = \sum_{n=0}^{\infty} \frac{(\alpha_1)_n ....(\alpha_p)_n}{n!(\beta_1)_n ...(\beta_q)_n} z^n, \tag{1.3.3}$$

where the *Pochhammer symbol* $(\alpha)_n$ is defined by

$$(\alpha)_n = \alpha(\alpha+1)\ldots(\alpha+n-1) = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)}, \quad (\alpha)_0 = 1.$$

We will be mostly interested in the hypergeometric functions which are solutions to the second order partial differential equation:

$$F''(z) + p(z)F'(z) + q(z)F(z) = 0 \tag{1.3.4}$$

for some holomorphic functions $p(z), q(z)$.

Definition 1.3.1. Let $z_1 \in \mathbb{C}$ be an isolated singularity for the holomorphic function $f(z)$. The singularity $z = z_1$ is called *regular* if there exists an exponent $\rho \in \mathbb{C}$ for which the function $(z - z_1)^{-\rho}F(z)$ is a holomorphic function in the neighborhood of z_1. The point ∞ is a regular singularity of the function $F(z)$ if $z = 0$ is a regular singularity of the function $F\left(\frac{1}{z}\right)$.

The following conditions on the coefficients $p(z)$ and $q(z)$ ensure that solutions to the equation (1.3.4) have only regular singularities (see [9]):

Theorem 1.3.2. (Fuchs) *Assume that equation (1.3.4) has solutions with singularities at the points $z_1, \ldots, z_n$ and ∞. Then these singularities are all regular if and only if the functions $p(z)$ and $q(z)$ have the form*

$$p(z) = \frac{p_0(z)}{(z - z_1)\ldots(z - z_n)} \tag{1.3.5}$$

and

$$q(z) = \frac{q_0(z)}{(z - z_1)^2\ldots(z - z_n)^2} \tag{1.3.6}$$

where $p_0(z)$ is a polynomial of degree $(n-1)$ and $q_0(z)$ is a polynomial of degree $2n-2$.

In the case that we have three regular (singular) points a, b, c, the differential equation (1.3.4) is called a *Riemann differential equation* and can be rewritten in the following form:

$$\frac{d^2w}{dz^2} + \left[\frac{1-\alpha-\alpha'}{z-a} + \frac{1-\beta-\beta'}{z-b} + \frac{1-\gamma-\gamma'}{z-c}\right]\frac{dw}{dz} + \\ + \left[\frac{\alpha\alpha'(a-b)(a-c)}{z-a} + \frac{\beta\beta'(b-c)(b-a)}{z-b} + \frac{\gamma\gamma'(c-a)(c-b)}{z-c}\right]\frac{w}{(z-a)(z-b)(z-c)} = 0.$$

The pairs of exponents with respect to the singular points $a; b; c$ are $\alpha, \alpha'; \beta, \beta'; \gamma, \gamma'$ respectively subject to the condition

$$\alpha + \alpha' + \beta + \beta' + \gamma + \gamma' = 1.$$

The complete set of solutions is denoted by the symbol

$$w = P\begin{Bmatrix} a & b & c & \\ \alpha & \beta & \gamma & z \\ \alpha' & \beta' & \gamma' & \end{Bmatrix} \tag{1.3.7}$$

1.3.1 Hypergeometric function

In the case when the singular points are $0, 1$ and ∞ we obtain the hypergeometric function ${}_2F_1(\alpha, \beta; \gamma; z)$:

$${}_2F_1(\alpha, \beta; \gamma; z) = P\begin{Bmatrix} 0 & \infty & 1 & \\ 0 & \alpha & 0 & z \\ 1-\gamma & \beta & \gamma-\alpha-\beta & \end{Bmatrix} \tag{1.3.8}$$

which is a solution to the *hypergeometric differential equation*

$$z(1-z)F''(z) + (\gamma - (1+\alpha+\beta)z)F'(z) - \alpha\beta F(z) = 0. \tag{1.3.9}$$

The exponents at $z = 0$ are $0, 1-\gamma$ and at $z = 1$ are $0, \gamma - \alpha - \beta$. The Taylor expansion for the hypergeometric function is

$${}_2F_1(\alpha, \beta; \gamma; z) = \sum_{n=0}^{\infty} \frac{(\alpha)_n(\beta)_n}{(\gamma)_n}\frac{z^n}{n!}. \tag{1.3.10}$$

Two linearly independent solution in the neighborhood of $z = 0$ are given by:

$$w_1 = {}_2F_1(\alpha, \beta; \gamma; z), \tag{1.3.11}$$

$$w_2 = z^{1-\gamma}{}_2F_1(\alpha-\gamma+1, \beta-\gamma+1; 2-\gamma; z), \tag{1.3.12}$$

and in the neighborhood of $z = 1$

$$w_1 = {}_2F_1(\alpha, \beta; \alpha+\beta+1-\gamma; 1-z), \tag{1.3.13}$$

$$w_2 = (1-z)^{\gamma-\alpha-\beta}{}_2F_1(\gamma-\beta, \gamma-\alpha, \gamma-\alpha-\beta+1, 1-z). \tag{1.3.14}$$

The derivative of the hypergeometric function is:

$${}_2F_1'(\alpha, \beta; \gamma; z) = \frac{\alpha\beta}{\gamma}{}_2F_1(\alpha+1, \beta+1; \gamma+1, z). \tag{1.3.15}$$

We will also need later increasing and decreasing solutions of the hypergeometric equation, which in the case $\alpha > 0,\ \beta > 0,\ \gamma > 0$ and $\gamma < \alpha + \beta + 1$ are given by:

$$\varphi^+(x) = {}_2F_1(\alpha, \beta; \gamma; z), \tag{1.3.16}$$

$$\varphi^-(x) = {}_2F_1(\alpha, \beta; \alpha+\beta+1-\gamma; 1-z). \tag{1.3.17}$$

1.3.2 Confluent hypergeometric function

The confluent hypergeometric function ${}_1F_1(a, b, z)$ (also denoted by $M(a, b, z)$ or $\Phi(a, b, z)$) can be obtained as the limit of Riemann's P-function

$$w = P \begin{Bmatrix} 0 & \infty & c & \\ 0 & -c & c+a & z \\ 1-b & 0 & -a & \end{Bmatrix} \tag{1.3.18}$$

provided $c \to \infty$.

Function $M(a, b, z)$ is a solution to the *Kummer differential equation*

$$zF''(z) + (b - z)F'(z) - aF(z) = 0 \tag{1.3.19}$$

The exponents at $z = 0$ are $0, 1 - b$.

The Taylor expansion for $M(a, b, z)$ is given by

$$M(a, b, z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \frac{z^n}{n!}. \tag{1.3.20}$$

Two linearly independent solutions to the Kummer differential equation are given by:

$$w_1 = M(a; b; z), \quad w_2 = z^{1-b} M(1 + a - b; 2 - b; z). \tag{1.3.21}$$

The increasing and decreasing solutions are given by:

$$\varphi^+(x) = M(a; b; z), \tag{1.3.22}$$
$$\varphi^-(x) = U(a, b, z) = \frac{\pi}{\sin(\pi b)} \left(\frac{M(a, b, z)}{\Gamma(1 + a - b)\Gamma(b)} - z^{1-b} \frac{M(1 + a - b, 2 - b, z)}{\Gamma(a)\Gamma(2 - b)} \right). \tag{1.3.23}$$

The asymptotics of M and U as $|z| \to \infty$ ($\Re z > 0$) is:

$$\varphi^+(z) = M(a, b, z) = \frac{\Gamma(b)}{\Gamma(a)} e^z z^{a-b} (1 + O(|z|^{-1})), \tag{1.3.24}$$
$$\varphi^-(z) = U(a, b, z) = z^{-a} (1 + O(|z|^{-1})), \tag{1.3.25}$$

and the derivative of the confluent hypergeometric function can be computed as

$$M'(a, b, z) = \frac{a}{b} M(a + 1, b + 1, z). \tag{1.3.26}$$

Chapter 2

Review of solvable models

In this chapter we present the necessary facts about OU (Ornstein-Uhlenbeck), CIR (Cox-Ingersoll-Ross) and Jacobi process and their lattice approximations given by birth and death processes (Charlier, Meixner and Hahn processes). These processes are chosen because they are associated with families of orthogonal polynomials, and are used as a starting point for constructing solvable Markov martingales in the next chapters. The main objective here is to express in each case the transition probability density $p(t, x_0, x_1)$ as an expansion in orthogonal polynomials and to study the boundary behavior of these processes.

In the first section we give a summary of results on orthogonal polynomials and a connection to the theory of Markov process (see [22]). In the second section we present a list of facts and formulas for the six chosen processes (these facts are used in chapters 3 and 5).

2.1 Introduction

Let X_t be a conservative Markov process on D. Assume that the process is positive recurrent, in which case $m(D)$ is finite, thus we can normalize it and assume that $m(dx)$ is a probability measure on D.

Consider the space $L_2(D, m)$ of real valued functions on D with inner product

$$(f, g)_m = \int_D f(x)g(x)m(dx). \tag{2.1.1}$$

One can check that the Markov generator $\mathcal{L}$ is symmetric in $L_2(D, m)$, thus its spectrum is real. Assume that $\mathcal{L}$ has a discrete spectrum $\{\lambda_n\}_{n\geq 0}$ and a set of eigenfunctions $\{\psi_n(x)\}_{n\geq 0}$, which form a complete orthogonal set in $L_2(D, m)$. The transitional probability density of X_t (with respect to $m(dx)$) is a kernel of the operator $e^{t\mathcal{L}}$, thus it is given by the following expansion:

$$p(t, x_0, x_1) = \sum_{n=0}^{\infty} \frac{e^{\lambda_n t}}{(\psi_n, \psi_n)_m} \psi_n(x_0)\psi_n(x_1). \tag{2.1.2}$$

We are interested in the case, when the eigenfunctions $\psi_n(x)$ are given by orthogonal polynomials.

2.1.1 Orthogonal polynomials

A system of polynomials $\{Q_n(x), n \in \mathcal{N},\ x \in D\}$, where deg$(Q_n(x)) = n$ and $\mathcal{N} = \mathbb{Z}^+ = \{0, 1, 2, \dots\}$ or $\mathcal{N} = \{0, 1, 2, \dots, N\}$, is called an *orthogonal system of polynomials* with respect to a (positive) measure $\mu(dx)$, if

$$(Q_n, Q_m)_\mu = \int_S Q_n(x)Q_m(x)\mu(dx) = d_n^2 \delta_{nm}, \quad n, m \in \mathcal{N}, \tag{2.1.3}$$

where S is the support of the measure $\mu(dx)$ (in our applications S is either the interval $D = (D^1, D^2)$ or a lattice when μ is a discrete measure).

The set $\{Q_n(x)\}$ of orthogonal polynomials can be obtained by a Gramm-Schmidt orthogonalization method starting with a (linearly independent) set of polynomials $\{1, x, x^2, \dots\}$.

It is a well known fact that orthogonal polynomials on the real line satisfy a *three term recurrence relation*

$$-\, xQ_n(x) = a_n Q_{n+1}(x) + b_n Q_n(x) + c_n Q_{n-1}(x), \quad n \geq 1, \tag{2.1.4}$$

where $b_n, c_n \neq 0$ and $c_n/b_{n-1} > 0$. (It was proved by Favard that the converse is true: a system of polynomials defined by the recurrence relation (2.1.4) is orthogonal with respect to some measure μ).

The main importance of the recurrence relation (2.1.4) for our applications lies in the fact that, if we know the coefficients a_n, b_n, c_n, then using "initial condition" $Q_{-1}(x) = 0$, $Q_0(x) = 1$ we can actually compute iteratively all orthogonal polynomials $Q_n(x)$.

Classical orthogonal polynomials are obtained as particular solutions to a differential equation of the form:

$$p_2(x)\frac{d^2 f(x)}{dx^2} + p_1(x)\frac{df(x)}{dx} - \lambda f(x) = 0, \tag{2.1.5}$$

where $p_2(x)$ and $p_1(x)$ are polynomials of degree at most 2 and 1. It can be shown that for

$$\lambda_n = np_1' + \frac{1}{2}n(n-1)p_2'', \tag{2.1.6}$$

equation (2.1.5) has a polynomial solution $f = q_n(x)$. These polynomials satisfy the orthogonality relation

$$\int_{D^1}^{D^2} q_n(x)q_m(x)\mu(dx) = d_n^2\delta_{nm},$$

where the measure $\mu(dx)$ has a density $\mu(dx) = m(x)dx$ which is a solution to the *Pearson equation*

$$(p_2(x)m(x))' = p_1(x)m(x),$$

and the endpoints D^1 and D^2 are zeros of $p_2(x)$ (or ∞). The system of orthogonal polynomials $q_n(x)$ is complete in the space $L_2(D, \mu)$.

Discrete orthogonal polynomials arise as solutions to the *hypergeometric type difference equation*

$$p_2(x)\Delta f(x) + p_1(x)\nabla^+ f(x) - \lambda f(x) = 0, \tag{2.1.7}$$

where

$$\Delta f(x) = f(x+1) - 2f(x) + f(x-1), \quad \nabla^+ f(x) = f(x+1) - f(x)$$

are the discrete analogs of $\frac{d^2f}{dx^2}$ and $\frac{df}{dx}$. It can be shown that under some technical conditions, for λ given by (2.1.6), equation (2.1.7) has a polynomial solution. These polynomials are orthogonal on the lattice $\{D^1, D^1+1, \ldots, D^2\}$ for some D^1 and D^2 and the orthogonality measure satisfies the discrete analog of the Pearson equation

$$\nabla^+(p_2(x)m(x)) = p_1(x)m(x).$$

2.1.2 Classification of classical orthogonal polynomials

Classical orthogonal polynomials of a continuous variable x are classified according to whether the polynomial $p_2(x)$ is constant, linear or quadratic:

(i) $\deg(p_2(x)) = 0$.

Let $p_2(x) = 1$ and $P_1(x) = -2x$. Then $D = (-\infty, \infty)$, $\lambda_n = -n$ and $m(x) = \exp(-x^2/2)/\sqrt{\pi}$ is a normal distribution. In this case we obtain the family of *Hermite polynomials* $H_n(x)$.

(ii) $\deg(p_2(x)) = 1$.

Let $p_2(x) = x$ and $p_1(x) = \alpha + 1 - x$. Then $D = (0, \infty)$, $\lambda_n = -n$ and $m(x) = x^\alpha e^{-x}/\Gamma(\alpha + 1)$ is the Gamma distribution. In this case we have the family of *Laguerre polynomials* $L_n^\alpha(x)$. These polynomials satisfy the orthogonality relation when $\alpha > -1$ (otherwise $m(D) = \infty$).

(iii) $\deg(p_2(x)) = 2$.

In this case it can be shown that one can have an infinite system of orthogonal polynomials if and only if $p_2(x)$ has exactly two zeros. Assume $p_2(x) = 1 - x^2$ and $p_1(x) = \beta - \alpha - (\alpha + \beta + 2)x$. Then $D = (-1, 1)$, $\lambda_n = -n(n + \alpha + \beta + 1)$ and $m(x) = (1 - x)^\alpha(1 + x)^\beta$ is the Beta kernel. The corresponding family of orthogonal polynomials is called *Jacobi polynomials* $P_n^{(\alpha,\beta)}(x)$ and it satisfies the orthogonality relation when $\alpha > -1, \beta > -1$ (otherwise $m(D) = \infty$).

The orthogonal polynomials of discrete variable are classified in exactly the same manner: if $\deg(p_2(x)) = 2$ and $p_2(x)$ has two zeros one obtains the most general *Hahn polynomials*, which are orthogonal with respect to the hypergeometric distribution. In the case $\deg(p_2(x)) = 1$ one can obtain *Meixner polynomials*, which are orthogonal with respect to the Pascal distribution and *Charlier polynomials*, which are orthogonal with respect to the Poisson distribution.

For a reference to orthogonal polynomials (as well as their q-deformed analogs) see [14].

2.1.3 Markov processes and orthogonal polynomials

Orthogonal polynomials enter the theory of Markov processes through the Markov generators. For example, the generator of the Jacobi process is given by

$$\mathcal{L} = (a - bx)\frac{d}{dx} + \frac{1}{2}\sigma^2 x(A - x)\frac{d^2}{dx^2},$$

and the eigenfunction equation

$$\mathcal{L}f(x) = \lambda f(x)$$

coincides with the (rescaled) hypergeometric equation of the form (2.1.5). Thus the eigenfunctions can be expressed through Jacobi polynomials, and the probability density admits an expansion in orthogonal polynomials given by equation (2.1.2).

It can be proved (see [20]) that Jacobi, CIR and OU processes are the only diffusion processes that are associated with orthogonal polynomials.

The discrete hypergeometric equation (2.1.7) corresponds to the generator of the birth and death process described in (1.2). Thus Charlier, Meixner and Hahn polynomials are associated with certain birth and death processes, which we will call by the name of corresponding system of orthogonal polynomials.

2.2 Solvable models

In this section we will present a list of facts about the processes discussed above, including: generator, speed measure and scale function (in case of diffusion process), spectrum and eigenfunctions of the generator, three term recurrence relation for eigenfunctions and the expansion of transitional probability density in the orthogonal basis. The processes are given in the following order: a diffusion process is followed by its lattice approximation given by a birth and death process.

The formulas concerning orthogonal polynomials, such as the differential equation, three term recurrence relation and orthogonal relation can be found in [14]. In each case the boundary behavior for diffusion processes was found using Feller's lemma (1.1.5).

2.2.1 Ornstein-Uhlenbeck process

- Generator:

$$\mathcal{L} = (a - bx)\frac{d}{dx} + \frac{1}{2}\sigma^2\frac{d^2}{dx^2}, \tag{2.2.1}$$

  where $b > 0$.

- Domain: $D = (-\infty, +\infty)$.

- Speed measure and scale function:

$$m(x) = \frac{1}{\sqrt{\pi\frac{\sigma^2}{b}}}\exp\left(-\frac{b}{\sigma^2}(x - \frac{a}{b})^2\right), \quad s'(x) = \exp\left(\frac{b}{\sigma^2}(x - \frac{a}{b})^2\right) \tag{2.2.2}$$

- Boundary conditions: Both $D^1 = -\infty$ and $D^2 = +\infty$ are natural boundaries for all choices of parameters.

- Probability function:

$$p^{(OU)}(t, x_0, x_1)m(x_1) = \frac{1}{\sqrt{\pi\frac{\sigma^2}{b}(1 - e^{-2bt})}}\exp\left(-\frac{\frac{b}{\sigma^2}(x_1 - x_0e^{-bt} - \frac{a}{b}(1 - e^{-bt}))^2}{(1 - e^{-2bt})}\right). \tag{2.2.3}$$

- Spectrum of the generator:

$$\lambda_n = -bn. \tag{2.2.4}$$

- Eigenfunctions of the generator:

$$\psi_n(x) = H_n\left(\sqrt{\frac{b}{\sigma^2}}(x - \frac{a}{b})\right), \tag{2.2.5}$$

where $H_n(x)$ are Hermite polynomials. The three term recurrence relation is:

$$H_{n+1} - 2xH_n(x) + 2nH_{n-1}(x) = 0. \tag{2.2.6}$$

- Orthogonality relation

$$\int_D \psi_n(x)\psi_m(x)m(x)dx = 2^n n!\delta_{nm}.$$

- Eigenfunction expansion of the probability function:

$$p^{(\mathrm{OU})}(t, x_0, x_1) = \sum_{n=0}^{\infty} \frac{e^{-bnt}}{2^n n!} H_n(y_0)H_n(y_1), \tag{2.2.7}$$

where $y_i = \sqrt{\frac{b}{\sigma^2}}(x_i - \frac{a}{b})$.

2.2.2 Charlier process

- Generator:

$$\mathcal{L} = (a - bx)\nabla_\delta^- + \frac{1}{2}\sigma^2\Delta_\delta, \tag{2.2.8}$$

where operators ∇_δ^- and Δ_δ are defined as:

$$\nabla_\delta^- f(x) = \frac{1}{\delta}(f(x) - f(x - \delta)), \quad \Delta_\delta f(x) = \frac{1}{\delta^2}(f(x + \delta) - 2f(x) + f(x - \delta)). \tag{2.2.9}$$

- Domain: $D = \{D^1, D^1 + \delta, D^1 + 2\delta, \ldots\}$, where $D^1 = D^1(\delta) = \frac{a}{b} - \frac{\sigma^2}{2b\delta}$. Notice that $D^1(\delta) \to -\infty$ as $\delta \searrow 0$.

- Spectrum of the generator:

$$\lambda_n = -bn. \tag{2.2.10}$$

- Eigenfunctions of the generator:

$$\psi_n(x) = C_n(y, \alpha), \tag{2.2.11}$$

where $\alpha = \frac{\sigma^2}{2b\delta^2}$, $y = (x - D^1)/\delta \in \{0, 1, 2, \dots\}$ and C_n are Charlier polynomials with the three term recurrence relation:

$$-yC_n(y, \alpha) = aC_{n+1}(y, \alpha) - (n + \alpha)C_n(y, \alpha) + nC_{n-1}(y, \alpha). \tag{2.2.12}$$

- Orthogonality measure (invariant measure) is given by the *Poisson distribution* $P(\alpha)$:

$$m(\{x\}) = e^{-\alpha}\frac{\alpha^y}{y!}, \quad y = (x - D^1)/\delta, \tag{2.2.13}$$

and orthogonality relation

$$\sum_{x \in D} \psi_n(x)\psi_m(x)m(\{x\}) = \alpha^{-n} n! \delta_{nm}.$$

- Eigenfunction expansion of the probability function:

$$p^{(\text{Charlier})}(t, x_0, x_1) = \sum_{n=0}^{\infty} e^{-bnt}\frac{\alpha^n}{n!}C_n(y_0, \alpha)C_n(y_1, \alpha), \tag{2.2.14}$$

where $y_i = (x_i - D^1)/\delta$.

2.2.3 CIR process

- Generator

$$\mathcal{L} = (a - bx)\frac{d}{dx} + \frac{1}{2}\sigma^2 x\frac{d^2}{dx^2}. \tag{2.2.15}$$

- Domain $D = [0, +\infty)$

- Speed measure and scale function:

$$m(x) = \frac{\theta^{\alpha+1}}{\Gamma(\alpha + 1)}x^\alpha e^{-\theta x}, \quad s'(x) = x^{-\alpha-1}e^{\theta x}, \tag{2.2.16}$$

where $\alpha = \frac{2a}{\sigma^2} - 1$ and $\theta = \frac{2b}{\sigma^2}$.

- Boundary conditions: $D^2 = +\infty$ is a natural boundary for all choices of parameters and

$$D^1 = \begin{cases} \text{exit, if } \alpha \le -1 \\ \text{regular, if } -1 < \alpha < 0 \\ \text{entrance, if } 0 \le \alpha \end{cases} \tag{2.2.17}$$

- Probability function:

$$p^{(CIR)}(t, x_0, x_1)m(x_1) = c_t \left(\frac{x_1 e^{bt}}{x_0}\right)^{\frac{1}{2}\alpha} \exp\left[-c_t(x_0 e^{-bt} + x_1)\right] I_\alpha\left(2c_t\sqrt{x_0 x_1 e^{-bt}}\right), \tag{2.2.18}$$

where $c_t \equiv -2b/(\sigma^2(e^{-bt} - 1))$ and I_α is the modified Bessel function of the first kind (see [8]).

- Spectrum of the generator:

$$\lambda_n = -bn. \tag{2.2.19}$$

- Eigenfunctions of the generator:

$$\psi_n(x) = L_n^\alpha(\theta x), \tag{2.2.20}$$

where $L_n^\alpha(y)$ are Laguerre polynomials of order α with the three term recurrence relation:

$$(n+1)L_{n+1}^\alpha(y) - (2n + \alpha + 1 - y)L_n^\alpha(y) + (n + \alpha)L_{n-1}^\alpha(y) = 0. \tag{2.2.21}$$

- Orthogonality relation:

$$\int_D \psi_n(x)\psi_m(x)m(x)dx = \frac{(\alpha+1)_n}{n!}\delta_{nm}.$$

- Eigenfunction expansion of the probability function:

$$p^{(\text{CIR})}(t, x_0, x_1) = \sum_{n=0}^{\infty} e^{-bnt} \frac{n!}{(\alpha+1)_n} L_n^\alpha(\theta x_0) L_n^\alpha(\theta x_1). \tag{2.2.22}$$

2.2.4 Meixner process

- Generator

$$\mathcal{L} = (a - bx)\nabla_\delta^+ + \frac{1}{2}\sigma^2 x \Delta_\delta. \tag{2.2.23}$$

- Domain $D = \{0, \delta, 2\delta, \dots\}$

- Spectrum of the generator:

$$\lambda_n = -bn. \tag{2.2.24}$$

- Eigenfunctions of the generator:

$$\psi_n(x) = M_n(x/\delta, \beta, c), \tag{2.2.25}$$

where $c = 1 - \frac{2b\delta}{\sigma^2}$ and $\beta = \frac{2a}{c\sigma^2}$ and $M_n(y, \beta, c)$ are Meixner polynomials with the three term recurrence relation:

$$(c-1)yM_n(y,\beta,c) = c(n+\beta)M_{n+1}(y,\beta,c) - (n+(n+\beta)c)M_n(y,\beta,c) + nM_{n-1}(y,\beta,c).$$

- Orthogonality measure (invariant measure) is given by the *negative binomial distribution* or *Pascal distribution* $Pa(\beta, c)$:

$$m(\{x\}) = \frac{(\beta)_y c^y}{y!(1-c)^\beta} \tag{2.2.26}$$

and orthogonality relation

$$\sum_{y=0}^{\infty} M_n(y,\beta,c)M_m(y,\beta,c)m(\{x\}) = \frac{c^{-n}n!}{(\beta)_n}\delta_{nm},$$

where $y = x/\delta$.

- Eigenfunction expansion of the probability function:

$$p^{(\text{Meixner})}(t, x_0, x_1) = \sum_{n=0}^{\infty} e^{-bnt}\frac{c^n(\beta)_n}{n!}M_n(y_0,\beta,c)M_n(y_1,\beta,c), \tag{2.2.27}$$

where $y_i = x_i/\delta \in \{0, 1, 2, \dots\}$.

2.2.5 Jacobi process

- Generator

$$\mathcal{L} = (a - bx)\frac{d}{dx} + \frac{1}{2}\sigma^2 x(A - x)\frac{d^2}{dx^2}. \tag{2.2.28}$$

- Domain $D = [0, A]$

- Speed measure and scale function:

$$m(x) = \frac{x^{\beta}(A-x)^{\alpha}}{A^{\alpha+\beta+1}B(\alpha+1,\beta+1)}, \quad s'(x) = x^{-\beta-1}(A-x)^{-\alpha-1}, \tag{2.2.29}$$

where $\alpha = \frac{2b}{\sigma^2} - \frac{2a}{\sigma^2 A} - 1$ and $\beta = \frac{2a}{\sigma^2 A} - 1$.

- Boundary behavior for the Jacobi process is the same as for CIR process at the left boundary:

$$D^1 = \begin{cases} \text{exit, if } \beta \leq -1 \\ \text{regular, if } -1 < \beta < 0 \\ \text{entrance, if } 0 \leq \beta \end{cases} \tag{2.2.30}$$

The same classification applies to right boundary, we only need to replace β by α. Notice that in the case when $a > 0$, $b > 0$ and $\frac{a}{b} < A$ (which means that mean-reverting level lies in the interval $(0, A)$), we have $\alpha > -1$ and $\beta > -1$ and thus both boundaries are not exit.

- Spectrum of the generator:

$$\lambda_n = -\frac{\sigma^2}{2} n(n - 1 + \frac{2b}{\sigma^2}). \tag{2.2.31}$$

- Eigenfunctions of the generator:

$$\psi_n(x) = P_n^{(\alpha,\beta)}(y), \tag{2.2.32}$$

where $y = (\frac{2x}{A} - 1)$ and $P_n^{(\alpha,\beta)}(y)$ are Jacobi polynomials with the three term recurrence relation:

$$\begin{aligned} yP_n^{(\alpha,\beta)}(y) = \; & \frac{2(n+1)(n+\alpha+\beta+1)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)} P_{n+1}^{(\alpha,\beta)}(y) + \\ & + \frac{\beta^2-\alpha^2}{(2n+\alpha+\beta)(2n+\alpha+\beta+2)} P_n^{(\alpha,\beta)}(y) + \\ & + \frac{2(n+\alpha)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)} P_{n-1}^{(\alpha,\beta)}(y). \end{aligned}$$

- Orthogonality relation

$$\int_D \psi_n(x)\psi_m(x)m(dx) = p_n^2 \delta_{nm} = \frac{(\alpha+1)_n(\beta+1)_n}{(\alpha+\beta+2)_{n-1}(2n+\alpha+\beta+1)n!}\delta_{nm}.$$

- Eigenfunction expansion of the probability function:

$$p^{(\text{Jacobi})}(t, x_0, x_1) = \sum_{n=0}^{\infty} \frac{e^{-\lambda_n t}}{p_n^2} P_n^{(\alpha,\beta)}(y_0) P_n^{(\alpha,\beta)}(y_1), \tag{2.2.33}$$

where $y_i = (\frac{2x_i}{A} - 1)$.

2.2.6 Hahn process

- Generator

$$\mathcal{L} = (a - bx)\nabla_\delta + \frac{\sigma^2}{2}x(A - x)\Delta_\delta. \tag{2.2.34}$$

- Domain $D = \{0, \delta, 2\delta, \dots, N\delta\}$

- Spectrum of the generator:

$$\lambda_n = -\frac{\sigma^2}{2}n(n - 1 + \frac{2b}{\sigma^2}). \tag{2.2.35}$$

- Eigenfunctions of the generator: Hahn polynomials

$$\psi_n(x) = Q_n(x/\delta; \alpha, \beta, N), \tag{2.2.36}$$

where the parameters α, β and N can be found from the following system of equations:

$$\begin{cases} \beta + \alpha + 1 = \frac{2b}{\sigma^2} - 1, \\ N + \beta + 1 = \frac{A}{\delta}, \\ N(\alpha + 1) = \frac{2a}{\sigma^2\delta}, \\ N \in \mathbb{N}. \end{cases} \tag{2.2.37}$$

Note that due to the restriction $N \in \mathbb{N}$, these equations do not have solutions for all values of parameters a, b, A, σ. In applications we start with parameters β, α, N and then compute a, b, A, σ.

Hahn polynomials satisfy the following three terms recurrence relation:

$$-yQ_n(y) = A_n Q_{n+1}(y) - (A_n + C_n)Q_n(y) + C_n Q_{n-1}(y), \tag{2.2.38}$$

where $Q_n(y) = Q_n(y; \alpha, \beta, N)$ and

$$\begin{cases} A_n = \frac{(n+\alpha+\beta+1)(n+\alpha+1)(N-n)}{(2n+\alpha+\beta+1)(2n+\alpha+\beta+2)}, \\ C_n = \frac{n(n+\alpha+\beta+N+1)(n+\beta)}{(2n+\alpha+\beta)(2n+\alpha+\beta+1)}. \end{cases}$$

- Orthogonality measure is given by the scaled *hypergeometric distribution*:

$$m(\{x\}) = \begin{pmatrix} \alpha + y \\ y \end{pmatrix} \begin{pmatrix} \beta + N - y \\ N - y \end{pmatrix} \tag{2.2.39}$$

Orthogonality relation:

$$\sum_{y=0}^{N} Q_n(y;\alpha,\beta,N)Q_m(y;\alpha,\beta,N)m(\{y\delta\}) = q_n^2\delta_{nm},$$
$$q_n^2 = \frac{(-1)^n(n+\alpha+\beta+1)_{N+1}(\beta+1)_n n!}{(2n+\alpha+\beta+1)(\alpha+1)_n(-N)_n N!}.$$

- Eigenfunction expansion of the probability function:

$$p^{(\text{Hahn})}(t,x_0,x_1) = \sum_{n=0}^{\infty} \frac{e^{-\lambda_n t}}{q_n^2} Q_n(y_0;\alpha,\beta,N)Q_n(y_1;\alpha,\beta,N), \tag{2.2.40}$$

where $y_i = x_i/\delta$.

Chapter 3

Stochastic Transformations

In this chapter we introduce the concept of *stochastic transformation* as a way to obtain new (stationary) solvable processes from the already known ones. We also require the new processes to be driftless (one can always introduce a drift later by a measure change or change of variables). The method consists of two ingredients: change of measure and change of variables. The main ideas and equations for finding stochastic transformations were introduced first in [2]. In the first section we state and prove the main theorem (3.2.1), which gives an explicit way to construct all stochastic transformations.

In the second section we study the properties of transformed processes and stochastic transformations and we give a criterion for recognizing when two processes are related by a stochastic transformation. We show how the well-known examples of CEV and quadratic volatility processes can be immediately obtained using stochastic transformations (from Bessel process and Brownian motion respectively) and we construct three new 6-7-parameter families of driftless processes, which can be reduced to OU, CIR or Jacobi processes. We show that these families exhibit different boundary behavior, namely the OU family has both natural boundaries, the CIR family has one natural and one killing (or exit) boundary and the Jacobi process has both killing (exit) boundaries.

In the last section we generalize this method to birth and death processes, which can be considered lattice approximations to the processes in continuous phase space.

3.1 Introduction

Let's ask a question: how can we transform a Markov diffusion process X_t in such a way that the transition probability density of the transformed process can be expressed through the probability density of X_t? First of all let's take a look at what transformations are available.

Given a stochastic process (X_t, P) one has the following obvious choice:

- *change of variables* (change of state space for the process):

$$X_t = (X_t, \mathrm{P}) \mapsto Y_t = (Y(X_t), \mathrm{P}),$$

  where $Y(x)$ is a diffeomorphism $Y : D_x \rightarrow D_y$.

- *change of measure*:

$$(X_t, \mathrm{P}) \mapsto (X_t, \mathrm{Q}),$$

  where measure Q is absolutely continuous with respect to P: $d\mathrm{Q}_t = Z_t d\mathrm{P}_t$.

- *time change*:

$$X_t = (X_t, \mathcal{F}_t, \mathrm{P}) \mapsto \tilde{X}_t = (X_{\tau_t}, \mathcal{F}_{\tau_t}, \mathrm{P}),$$

  where τ_t is an increasing stochastic process, $\tau_0 = 0$.

- and at last we can combine the above transformations in any order.

In this chapter we will discuss only the first two types of transformations, while the time change will be discussed in detail in chapter 5.

The next obvious question is: do these transformations preserve the solvability of the process? The answer is always "yes" for the change of variables transformation (we assume that the function $Y(x)$ is invertible): the probability density of the process $Y_t = Y(X_t)$ is given by:

$$p_Y(t, y_0, y_1) = p_X(t, X(y_0), X(y_1)), \quad y_i \in D_y,$$

where $X = X(y) = Y^{-1}(y)$ and the speed measure of Y_t is

$$m_Y(dy) = m_Y(y)dy = m_X(X(y))X'(y)dy.$$

What can we say about the measure change transformation? We want the new measure Q to be absolutely continuous with respect to P, thus there exists a positive process Z_t, such that

$$\frac{d\mathrm{Q}_t}{d\mathrm{P}_t} = Z_t.$$

The process Z_t must be a (local) martingale. Under what conditions on Z_t does this transformation preserves solvability?

Informally speaking, the probability density of the process $X_t^{\mathrm{Q}} = (X_t, \mathrm{Q})$ is given by

$$p_{X^{\mathrm{Q}}}(t, x_0, x_1) m_{X^{\mathrm{Q}}}(x_1) = E^{\mathrm{Q}}(\delta(X_{s+t}^{\mathrm{Q}} - x_1) | X_s = x_0),$$

where we used the fact that the transformed process is stationary. Using a formula of change of measure under conditional expectation we obtain

$$p_{X^{\mathrm{Q}}}(t, x_0, x_1) m_{X^{\mathrm{Q}}}(x_1) = \frac{1}{Z_s} E^{\mathrm{P}}(Z_{s+t}\delta(X_{s+t}^{\mathrm{Q}} - x_1) | X_s = x_0).$$

Since we want the transformed process to be Markov, we see that this can be the case if and only if Z_s depends only on the value of X_s, thus the process Z_t can be represented as

$$Z_t = h(X_t, t),$$

for some positive function $h(x, t)$.

The next step is to note that the dynamics of X_t under the new measure is:

$$dX_t = \left(b(X_t) + \sigma^2(X_t) \frac{h_x(X_t, t)}{h(X_t, t)} \right) dt + \sigma(X_t) dW_t^{\mathrm{Q}}.$$

Now we see that if we want the transformed process to be stationary, h_x/h must be independent of t, which gives us the following expression for the function $h(x, t)$:

$$h(x, t) = h(x) g(t).$$

Since we don't want to introduce killing in the interior of the domain D, the process $Z_t = h(X_t)g(t)$ must be a martingale (at least a local one), thus we have the following equation

$$\frac{dg(t)}{dt} h(x) + g(t) \mathcal{L}_X h(x) = 0.$$

After separating the variables we find that there exists a constant ρ such that $g(t) = e^{-\rho t}$, and the function $h(x)$ is a solution to the following "eigenfunction equation":

$$\mathcal{L}_X h(x) = \rho h(x). \tag{3.1.1}$$

This discussion leads us to our main definition:

Definition 3.1.1. The *stochastic transformation* is a triple

$$\{\rho, h, Y\} \tag{3.1.2}$$

where ρ and $h(x)$ define the absolutely continuous measure change through the formula

$$d\mathrm{P}_t^h = \exp(-\rho t)h(X_t)d\mathrm{P}_t, \tag{3.1.3}$$

and $Y(x)$ is a diffeomorphism $Y : D_x \to D_y \subseteq \mathbb{R}$, such that the process $(Y_t, \mathrm{P}^h) = (Y(X_t), \mathrm{P}^h)$ is a driftless process.

Remark 3.1.2. We deliberately do not require Y_t to be a (local) martingale, since as we will see later Y_t is not conservative in general. Though using lemma (1.1.3) we see that Y_t is a driftless process if and only if Y_t stopped at the boundary of D_y is a (local) martingale.

3.2 Main Theorem

The following theorem gives an explicit way to find all stochastic transformations defined in (3.1.1):

Theorem 3.2.1. *Let X_t be a stationary Markov diffusion process under the measure P on the domain $D_x \subseteq \mathbb{R}$ and admitting a Markov generator $\mathcal{L}_X$ of the form*

$$\mathcal{L}_X f(x) = b(x)\frac{df(x)}{dx} + \frac{1}{2}\sigma^2(x)\frac{d^2 f(x)}{dx^2}. \tag{3.2.1}$$

We assume that both boundaries of D_x are inaccessible. Assume $\rho \geq 0$.

Then $\{\rho, h, Y\}$ is a stochastic transformation if and only if

$$\begin{cases} h(x) = c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x), \\ Y(x) = \frac{1}{h(x)}(c_3\varphi_\rho^+(x) + c_4\varphi_\rho^-(x)), \end{cases} \tag{3.2.2}$$

where $c_i \in \mathbb{R}$ are parameters ,$c_1, c_2 \geq 0$, $c_1c_4 - c_2c_3 \neq 0$, and $\varphi_\rho^+(x)$ and $\varphi_\rho^-(x)$ are increasing and decreasing solutions, respectively, to the differential equation

$$\mathcal{L}_X\varphi(x) = \rho\varphi(x).$$

Before we give a proof of this theorem we need to present some theory.

3.2.1 Doob's h-transform

Definition 3.2.2. A positive function $h(x)$ is called *ρ-excessive* for the process X if the following two statements hold true:

(i) $e^{-\rho t}E(h(X_t)|X_0 = x) \leq h(x)$

(ii) $\lim\limits_{t\to 0} E(h(X_t)|X_0 = x) = h(x)$.

An ρ-excessive function h is called *ρ-invariant* if for all $x \in D_x$ and $t \geq 0$

$$e^{-\rho t}E(h(X_t)|X_0 = x) = h(x).$$

Remark 3.2.3. Function $h(x)$ is ρ-excessive (ρ-invariant) if and only if the process $\exp(-\rho t)h(X_t)$ is a positive supermartingale (martingale). Thus, if a function $h(x)$ is zero at some x_0, then $h \equiv 0$ in D_x.

One can check that for every $x_1 \in D_x$ functions $\varphi_\rho^+(x)$, $\varphi_\rho^-(x)$ and $x \to G_X(\rho, x, x_1)$ are ρ-excessive. As the following lemma shows, these functions are *minimal* in the sense that any other excessive function (except the trivial example $h \equiv C$) can be expressed as a linear combination of them (see [4]):

Lemma 3.2.4. *Let $h(x)$ be a positive function on D_x such that $h(x_0) = 1$. Then h is ρ-excessive if and only if there exists a probability measure ν on $D_x = [D^1, D^2]$, such that for all $x \in D_x$*

$$\begin{aligned} h(x) &= \int_{[D^1,D^2]} \frac{G_X(\rho,x,x_1)}{G_X(\rho,x_0,x_1)}\nu(dx_1) = \\ &= \int_{(D^1,D^2)} \frac{G_X(\rho,x,x_1)}{G_X(\rho,x_0,x_1)}\nu(dx_1) + \frac{\varphi_\rho^-(x)}{\varphi_\rho^-(x_0)}\nu(\{D^1\}) + \frac{\varphi_\rho^+(x)}{\varphi_\rho^+(x_0)}\nu(\{D^2\}). \end{aligned}$$

Measure ν is called the representing measure of h.

Definition 3.2.5. The coordinate process X_t under the measure P^h defined by

$$\mathrm{P}^h(A|X_0 = x) = E\left(e^{-\rho t}\frac{h(X_t)}{h(x)}\mathbb{I}_A|X_0 = x\right), \tag{3.2.3}$$

is called *Doob's h-transform* or *ρ-excessive transform* of X. We will denote this process by (X, P^h) (or in short X^h).

The following lemma gives the expression of the main characteristics of the h-transform of X:

Lemma 3.2.6. *The process X^h is a regular diffusion process with:*

- *Generator*

$$\mathcal{L}_{X^h} = \frac{1}{h}\mathcal{L}_X h - \rho \tag{3.2.4}$$

with drift and diffusion terms given by

$$b_{X^h}(x) = b_X(x) + \sigma_X^2(x)\frac{h_x(x)}{h(x)}, \quad \sigma_{X^h}(x) = \sigma_X(x). \tag{3.2.5}$$

- *The speed measure and the scale function of the process X^h are*

$$m_{X^h}(x) = h^2(x)m_X(x), \quad s'_{X^h}(x) = h^{-2}(x)s'(x). \tag{3.2.6}$$

- *The transition probability density (with respect to the speed measure $m_{X^h}(dx_1)$)*

$$p_{X^h}(t, x_0, x_1) = \frac{e^{-\rho t}}{h(x_0)h(x_1)}p_X(t, x_0, x_1). \tag{3.2.7}$$

- *The Green function*

$$G_{X^h}(\lambda, x_0, x_1) = \frac{1}{h(x_0)h(x_1)}G_X(\rho + \lambda, x_0, x_1). \tag{3.2.8}$$

- *The killing measure*

$$k_{X^h}(dx) = \frac{m_X(x)\nu(dx)}{G_{X^h}(0, X_0, x)}. \tag{3.2.9}$$

Proof. We will briefly sketch the proof. Using the formula (3.2.3) we find the semigroup for the process X^h:

$$P_{X^h}(t)f(x) = e^{-\rho t}\frac{1}{h(x)}P_X(t)(hf)(x),$$

and from this formula we find the expression for the Markov generator and formulas for the drift and diffusion. Now,

$$B(x) = \int^x \frac{2b_{X^h}(y)}{\sigma_{X^h}^2(y)}dy = \int^x \frac{2b_X(y)}{\sigma_X^2(y)}dy + 2\log(h(x)),$$

from which formula we compute the expressions for the speed measure, scale function, probability density and Green function.

Finally, let's prove the formula for the killing measure: the infinitesimal killing rate c_{X^h} can be computed as $\mathcal{L}_{X^h}1$, thus we find:

$$\mathcal{L}_{X^h}1 = \frac{1}{h}\mathcal{L}_X h - \rho = \frac{1}{h}\mathcal{L}_X \int_{[D^1,D^2]} \frac{G_X(\rho,x,x_1)}{G_X(\rho,X_0,x_1)}\nu(dx_1) - \rho.$$

Now assuming that $\nu(dx) = \nu(x)dx$ and using the fact that the Green function satisfies the following equation

$$\mathcal{L}_X G_X(\rho,x,x_1) = \rho G_X(\rho,x,x_1) + \delta(x_1 - x),$$

we find

$$c_{X^h}(x) = \mathcal{L}_{X^h}1 = \frac{1}{h}\int_{[D^1,D^2]} \frac{\rho G_X(\rho,x,x_1) + \delta(x_1 - x)}{G_X(\rho,X_0,x_1)}\nu(x_1)dx_1 - \rho = \frac{\nu(x)}{h(x)G_X(\rho,X_0,x)},$$

thus the killing measure can be computed as

$$k_{X^h}(dx) = c_{X^h}(x)m_{X^h}(x)dx = \frac{m_X(x)\nu(dx)}{G_{X^h}(0,X_0,x)},$$

which ends the proof. □

Remark 3.2.7. Note that we have a nonzero killing measure on the boundary of D_x and no killing in the interior of D_x if and only if the representing measure $\nu(dx)$ is supported on the boundary of D_x, that is

$$h(x) = \nu(\{D^1\})\varphi_\rho^-(x) + \nu(\{D^2\})\varphi_\rho^+(x) = c_1\varphi_\rho^-(x) + c_2\varphi_\rho^+(x).$$

Remark 3.2.8. To understand the probabilistic meaning of the Doob's h-transform it is useful to consider the procedure of constructing new process by conditioning X_t on some event (the event A can be that the process stays in some interval or that it has some particular maximum or minimum value). The mathematical description follows:

Let the probability density $p_t(x,y|A(t_0,t_1))$ be the conditional density defined as follows:

$$p_t(x,y|A(t_0,t_1))dy = \mathrm{P}(X_{t+\delta t} \in dy | X_t = x, A(t_0,t_1)).$$

Assume that the event $A(t_0,t_1)$ is $\mathcal{F}_{t_1}$-measurable and it satisfies the semigroup property: $\mathrm{P}(A(t_0,t_1)|B) = \mathrm{P}(A(t_0,t') \cap A(t',t_1)|B)$ for any event B and $t' \in (t_0,t_1)$.

Let the probability $\pi(x,t;A(t_0,t_1)$ be defined as

$$\pi(x,t;A(t_0,t_1)) = \mathrm{P}(A(t_0,t_1)|X(t) = x).$$

Then function π satisfies the following backward Kolmogorov equation:

$$\frac{\partial \pi}{\partial t} + \mathcal{L}_X \pi = \frac{\partial \pi(x,t;A(t,T))}{\partial t} + \frac{1}{2}\sigma^2(x)\frac{\partial^2 \pi(x,t;A(t,T))}{\partial x^2} + b(x)\frac{\partial \pi(x,t;A(t,T))}{\partial x} = 0.$$

The boundary conditions depend on event $A(t,T)$.

The conditioned drift $b(x;A(t,T)$ and conditioned volatility $\sigma(x;A(t,T)$ are given by

$$\begin{aligned} b(x;A(t,T)) &= b(x) + \sigma^2(x)\frac{\pi_x(x,t;A(t,T))}{\pi(x,t;A(t,T))}, \\ \sigma(x;A(t,T)) &= \sigma(x). \end{aligned} \tag{3.2.10}$$

Lets consider some examples of the h-transform:

(i) **Brownian bridge through h-transform** Let $X_t = W_t$ be the brownian motion. Fix some x_0 and consider the event $A(t;T) = \{X_T = x_1\}$. Then the function $\pi(x,t;A(t,T))$ is the probability density of Brownian motion:

$$\pi(x,t;A(t,T)) = p_{T-t}(x,x_1) = \frac{1}{\sqrt{2\pi(T-t)}}\exp\left(-\frac{(x_1-x)^2}{2(T-t)}\right).$$

One can show that X^h is just the Brownian bridge - brownian motion conditioned on the event $X_T = x_1$ (note that this process is not time-homogeneous since in this case we are not using the ρ-excessive transform).

The conditional drift of the process X^h, computed by the formula (3.2.10) is equal to

$$b(x;A(t,T)) = \frac{\pi_x}{\pi} = \frac{x_0 - x}{T-t},$$

which is another way to prove that X^h is a Brownian bridge.

(ii) **Brownian Motion, $X_t = W_t$.** Increasing and decreasing solutions to equation

$$\mathcal{L}^X f = \frac{1}{2}\frac{d^2 f(x)}{dx^2} = \rho f(x)$$

are given by

$$\varphi_\rho^+(x) = e^{\sqrt{2\rho}x}, \qquad \varphi_\rho^-(x) = e^{-\sqrt{2\rho}x}$$

Note that in this case the process $e^{-\rho t}h(X_t)$ is a martingale. Let $h(x) = \varphi_\rho^+(x)$, then the process $X_t^h = W_t^h + \sqrt{2\rho}t$ is Brownian motion with drift. Note that since $e^{-\rho t}h(X_t)$ is a martingale, the transformed process is still conservative, but it has a completely different behavior: for example the $X_t^h \to +\infty$ as $t \to \infty$ and $D^2 = \infty$ becomes an attractive boundary. This is a typical situation for the ρ-excessive transforms: as we will see later, the following result is true: the transformed process X^h is either nonconservative with killing at the boundary, or in the case it is conservative, X_t^h converges to one of the two boundary points as $t \to \infty$.

We will return to the example of Brownian motion later in section (3.4.1).

(iii) Let X_t be a transient process with zero killing measure. Then $X_t \to D^1$ or $X_t \to D^2$ with probability 1 as $t \to \zeta$. Let X^+ be the $\varphi_0^+(x)$ transform of X. Note that $\varphi_0^+(x)$ is a constant multiple of $\mathrm{P}_x(X_t \to D^2$ as $t \to \zeta)$, thus X^+ is identical in law to X, given that $X_t \to D^2$ as $t \to \zeta$, or otherwise X^+ has the property:

$$\mathrm{P}_x(\lim_{t \to \zeta} X_t^+ = D^2) = 1.$$

3.2.2 Proof of the main theorem

Now we are ready to give the proof of the main theorem (3.2.1):

Proof. Lets prove first that function $h(x) = c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x)$ can be used to define a measure change. Define the stopping times

$$\tau_n = H_{D^1+1/n} \wedge H_{D^2-1/n} = \inf\{s : \; X_s \notin (D^1 + 1/n, D^2 - 1/n)\}.$$

As we proved in (1.1.8), the process $\exp(-\rho t)h(X_t)$ stopped at τ_n is a martingale for every n. Since the boundaries are inaccessible we have

$$\tau_n \nearrow +\infty, \quad \text{as } n \to \infty,$$

thus the process $\exp(-\rho t)h(X_t)$ is a local martingale. Since it is also positive it is actually a supermartingale (by Fatou lemma). Thus

$$E(e^{-\rho t}h(X_t)|X_0 = x) \leq h(x), \tag{3.2.11}$$

and $\exp(-\rho t)h(X_t)$ correctly defines an absolutely continuous measure change.

Lemmas (3.2.4) and (3.2.6) show that the converse statement is also true: if function $h(x)$ can be used to define an absolutely continuous measure change, then $e^{\rho t}h(X_t)$ is a supermartingale, thus $h(x)$ is a ρ-excessive function. Since we want the transformed process X^h to have no killing in the interior of D_x, representing measure $\nu(dx)$ must be supported at the boundaries (see remark (3.2.7)), thus we have the representation $h(x) = c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x)$.

To prove the second statement of the theorem, we note first that $Y(X_t)$ has a zero drift if and only if it is in the natural scale: thus we need to check that function $Y(x)$ given by formula (3.2.2) is equal to the scale function $s_{X^h}(x)$ (up to affine transformation), which we can check by direct computation

$$Y'(x) = \frac{d}{dx}\left(\frac{\varphi(x)}{h(x)}\right) = \frac{\varphi'(x)h(x) - h'(x)\varphi(x)}{h^2(x)} = \frac{W_{\varphi,h}(x)}{h^2(x)} = s'_{X^h}(x),$$

where φ is an arbitrary solution of $\mathcal{L}_X\varphi = \rho\varphi$ linearly independent of h, (thus it can be represented as $\varphi = c_3\varphi_\rho^+ + c_4\varphi_\rho^-$ with $c_1c_4 - c_2c_3 \neq 0$). This ends the proof of the main theorem. □

Remark 3.2.9. The statement that $Y(X_t)$ is a driftless process is an analog of the following lemma:

Lemma 3.2.10. *Let* $\mathrm{Q} \cong \mathrm{P}$*, and* $Z_t = \frac{d\mathrm{Q}_t}{d\mathrm{P}_t}$*. An adapted cadlag process* M_t *is a* P*-local martingale if and only if* M_t/Z_t *is a* Q*-local martingale.*

Note that this lemma does not assume that the process is a diffusion process and one could use it to define stochastic transformations for arbitrary Markov processes. One could argue as follows: $e^{-\rho t}h(X_t)$ and $e^{-\rho t}(c_3\varphi_\rho^+(X_t) + c_4\varphi_\rho^-(X_t))$ are local martingales under P, thus if we define a measure change density $Z_t = e^{-\rho t}h(X_t)$ the process

$$Y(X_t) = \frac{e^{-\rho t}(c_3\varphi_\rho^+(X_t) + c_4\varphi_\rho^-(X_t))}{Z_t} = \frac{c_3\varphi_\rho^+(X_t) + c_4\varphi_\rho^-(X_t)}{h(X_t)}$$

is a Q local martingale. However, as we will see later, in some cases Z_s is not a martingale, thus the measure change is not equivalent in general and we can't use this argument.

3.2.3 Generalization of the method of stochastic transformations to arbitrary multidimensional Markov processes

The method of constructing stochastic transformations described in theorem 3.2.1 can be generalized to jump processes and multidimensional Markov processes: let X_t be a stationary process on the domain

$D_x \subset \mathbb{R}^d$ with Markov generator $\mathcal{L}_X$. The first step is to find two linearly independent solutions to the equation

$$\mathcal{L}_X \varphi = \rho\varphi.$$

Assume that for some choice of c_1, c_2 the function $h(x) = c_1\varphi_1(x) + c_2\varphi_2(x)$ is positive. Then one has to prove that the process

$$Z_t = e^{-\rho t} h(X_t)$$

is a local martingale (or a supermartingale), thus it can be used to define a new measure P^h by the formula (3.2.3). Then we define the function $Y(x) : \mathbb{R}^d \to \mathbb{R}$ as

$$Y(x) = \frac{c_3\varphi_1(x) + c_4\varphi_2(x)}{h(x)}.$$

Now one can check that the generator of X_t^h is given by

$$\mathcal{L}_{X^h} = \frac{1}{h}\mathcal{L}_X h - \rho,$$

and to prove that the (one-dimensional) process $Y(X_t)$ is driftless one could argue as follows:

$$\mathcal{L}_{X^h} Y(x) = \left(\frac{1}{h}\mathcal{L}_X h - \rho\right)\frac{\varphi}{h} = \frac{1}{h}\mathcal{L}_X\left(h\frac{\varphi}{h}\right) - \rho\frac{\varphi}{h} = \frac{1}{h}\mathcal{L}_X\varphi - \rho\frac{\varphi}{h} = 0,$$

since $\varphi(x) = c_3\varphi_1(x) + c_4\varphi_2(x)$ is also a solution to $\mathcal{L}_X\varphi = \rho\varphi$.

Note that this "proof" does not use any information about the process X_t and thus it is very general. We will see later in section 3.5 how this method leads to a natural generalization of birth and death processes.

3.3 Properties of stochastic transformations

The following lemma summarizes the main characteristics of the process Y_t:

Lemma 3.3.1. *The process $Y_t = (Y(X_t), \mathrm{P}^h)$ is a regular diffusion process with*

- *Generator*

$$\mathcal{L}_{X^h} = \frac{1}{2}\sigma_Y^2(y)\frac{d^2}{dy^2} \tag{3.3.1}$$

where volatility function is given by:

$$\sigma_Y(Y(x)) = \sigma_X(x)Y'(x) = \sigma_X(x)\frac{CW(x)}{h^2(x)}. \tag{3.3.2}$$

($W(x) = s'_X(x)$ is the Wronskian of $\varphi_\rho^+, \varphi_\rho^-$).

- *The speed measure and the scale function of the process Y_t are*

$$m_Y(y) = \frac{1}{\sigma_Y^2(y)}, \quad s_Y(y) = y. \tag{3.3.3}$$

- *The transition probability density (with respect to the speed measure $m_Y(dy_1)$) is*

$$p_Y(t, y_0, y_1) = p_{X^h}(t, x_0, x_1) = \frac{e^{-\rho t}}{h(x_0)h(x_1)} p_X(t, x_0, x_1), \tag{3.3.4}$$

where $y_i = Y(x_i)$.

- *The Green function*

$$G_Y(\lambda, y_0, y_1) = G_{X^h}(\lambda, x_0, x_1) = \frac{1}{h(x_0)h(x_1)} G_X(\rho + \lambda, x_0, x_1). \tag{3.3.5}$$

Lemma 3.3.2. *If $c_2 = 0$ $(c_1 = 0)$ the domain D_y of the process Y_t is an interval of the form $[y_0, \infty)$ or $(-\infty, y_0]$, where $y_0 = c_3/c_1$ $(y_0 = c_4/c_2)$. In the case $c_1 \neq 0$ and $c_2 \neq 0$, the domain D_y is a bounded interval of the form $[y_0, y_1] = [c_4/c_2, c_1/c_3]$ or $[y_0, y_1] = [c_1/c_3, c_4/c_2]$.*

Proof. Remember that

$$Y(x) = \frac{c_3 \varphi_\rho^+(x) + c_4 \varphi_\rho^-(x)}{c_1 \varphi_\rho^+(x) + c_2 \varphi_\rho^-(x)}$$

Since both boundaries are inaccessible , it is true that one of the φ_ρ^+, φ_ρ^- is bounded and the other is unbounded near each boundary, thus

$$D_y^1 = \lim_{x \searrow D_x^1} Y(x) = \frac{c_4}{c_2}, \quad D_y^2 = \lim_{x \nearrow D_x^2} Y(x) = \frac{c_1}{c_3}.$$

□

Lemma 3.3.3. *The processes X^h and Y are transient.*

Proof. This follows from the fact that

$$\lim_{\lambda \searrow 0} G_{X^h}(\lambda, x_0, x_1) = \frac{1}{h(x_0)h(x_1)} G_X(\rho, x_0, x_1) < \infty,$$

thus X_t^h is a transient process (see chapter 1). □

The previous lemma tells us that with probability one X_t^h (and Y_t) visits every point in its domain only a finite number of times. Thus it converges as $t \to \infty$. Since with probability one it can not converge to a point in the interior of D_x, it must converge to the point on the boundary (or to the cemetery point Δ_∞ if the process is not conservative). Actually an even stronger result can be obtained by means of martingale theory:

Lemma 3.3.4. *If the process Y_t is conservative, then*

$$\mathrm{P}^h(\lim_{t\to\infty} Y_t = Y_\infty) = 1. \tag{3.3.6}$$

and Y_∞ is integrable. The random variable Y_∞ is supported at the boundary of the interval D_y: in the case $D_y = [y_1, \infty)$ we have $Y_\infty = y_1$ a.s., while in the general case $D_y = [y_1, y_2]$ we have that Y_t converges to Y_∞ also in L_1 and the distribution of Y_∞ is

$$\mathrm{P}(Y_\infty = y_2 | Y_0 = y_0) = \frac{y_0 - y_1}{y_2 - y_1}, \quad \mathrm{P}(Y_\infty = y_1 | Y_0 = y_0) = \frac{y_2 - y_0}{y_2 - y_1}. \tag{3.3.7}$$

Proof. If Y_t is conservative, then Y_t is a local martingale bounded from below (or above), thus it is a supermartingale bounded from below (or a submartingale bounded from above), thus it converges as $t \to \infty$. If $D_y = [y_1, y_2]$ and the process Y_t is bounded, then it is a uniformly integrable martingale and it converges to Y_∞ also in L_1, moreover

$$Y_t = E(Y_\infty | \mathcal{F}_t),$$

thus $y_0 = EY_t = EY_\infty$, from which we can find the distribution of Y_∞. □

Lemma 3.3.5. *The stochastic transformation $(X, \mathrm{P}) \mapsto (Y, \mathrm{P}^h)$ given by $\{\rho, h(x), Y(x)\}$ is invertible. The inverse transformation $(Y, \mathrm{P}^h) \mapsto (X, \mathrm{P})$ is given by $\{-\rho, 1/h(x), X(y)\}$.*

Proof. If the process X_t is conservative with inaccessible boundaries, the process Y_t will be a transient driftless process (possibly with accessible boundaries). Let's check that the process

$$Z_t = e^{\rho t} \frac{1}{h(X_t^h)}$$

is a P^h martingale:

$$E_x^{\mathrm{P}^h}(Z_t) = E_x^{\mathrm{P}} \left(e^{-\rho t} \frac{h(X_t^h)}{h(x)} Z_t \right) = \frac{1}{h(x)} E_x^{\mathrm{P}} 1 = Z_0,$$

since the initial process X_t is conservative and $E_x^{\mathrm{P}} 1 = 1$. Thus the transformation $\{-\rho, 1/h(x), X(y)\}$ maps the process $Y_t = (Y(X_t), \mathrm{P}^h)$ back into the process $X_t = (X_t, \mathrm{P})$. □

Definition 3.3.6. We will say that (X, P) and (Y, Q) *are related by a stochastic transformation* and will denote it by writing

$$X \sim Y$$

if there exists a stochastic transformation $\{\rho, h, Y\}$ which maps $(X, \mathrm{P}) \mapsto (Y, \mathrm{P}^h)$.

Lemma 3.3.7. *The relation $X \sim Y$ is an equivalence relation.*

Proof. We need to check that for all X, Y and Z

(i) $X \sim X$

(ii) if $X \sim Y$, then $Y \sim X$

(iii) if $X \sim Y$ and $Y \sim Z$, then $X \sim Z$.

The first property is obvious; the second was proved in the previous lemma. To check the third, let $\{\rho_1, h_1(x), Y(x)\}$ ($\{\rho_2, h_2(y), Z(y)\}$) be the stochastic transformation relating X and Y (Y and Z). Then $\{\rho_1 + \rho_2, h_1(x)h_2(Y(x)), Z(Y(x))\}$ is a stochastic transformation mapping $X \mapsto Z$. □

Thus "$\sim$" relation divides all the Markov stationary driftless diffusions into equivalence classes, which will be denoted by

$$\mathfrak{M}(X) = \mathfrak{M}(X, \mathrm{P}) = \{(Y, \mathrm{Q}) : \quad (Y, \mathrm{Q}) \sim (X, \mathrm{P})\}. \tag{3.3.8}$$

The following lemma gives a convenient criterion to check whether two processes are related by a stochastic transformation:

Lemma 3.3.8. *Let X_t be a diffusion process with Markov generator*

$$\mathcal{L}_X f(x) = b(x)\frac{df(x)}{dx} + \frac{1}{2}\sigma^2(x)\frac{d^2 f(x)}{dx^2}.$$

With this diffusion we will associate a function $J_X(z) = I_X(x(z))$, where

$$I_X(x) = \frac{1}{4}\left(\sigma(x)\sigma''(x) - \frac{1}{2}(\sigma'(x))^2 + 2\left[\frac{2b(x)\sigma'(x)}{\sigma(x)} - b'(x) - \frac{b^2(x)}{\sigma^2(x)}\right]\right). \tag{3.3.9}$$

and the change of variables $x(z)$ is defined through $x'(z) = \sigma(x)$. Then the function $I_X(z)$ is an invariant of the diffusion X_t in the following sense:

(i) If $Y_t = (Y(X_t), \mathrm{P})$, then $J_Y(z) = J_X(z)$ for all z.

(ii) If $X_t^h = (X_t, \mathrm{P}^h)$ is an ρ-excessive transform of X_t, then $J_{X^h}(z) = J_X(z) - \rho$.

(iii) If $X \sim Y$, then $J_Y(z) = J_X(z) - \rho$, where ρ is the same as in stochastic transformation $\{\rho, h, Y\}$ which relates X and Y.

We prove this lemma in the chapter 4 after we develop the necessary tools for this.

Example 3.3.9. Let's use lemma 3.3.8 to check that the *quadratic volatility family* with volatility functions of the form

$$\sigma_Y(y) = a_2 y^2 + a_1 y + a_0 \tag{3.3.10}$$

is related to a Brownian motion $X_t = W_t$ by a stochastic transformation. It is easy to see that $J_W \equiv 0$ and $J_Y \equiv const$, thus $Y_t \sim W_t$.

Example 3.3.10. As another application of lemma 3.3.8 let's show that a Bessel process

$$dX_t = adt + \sqrt{X_t} dW_t$$

is related to CEV processes (constant-elasticity-of-variance)

$$dY_t = Y_t^\theta dW_t.$$

Function $I_X(x)$ is given by

$$I_X(x) = \frac{1}{4}\left(-\frac{1}{4x} - \frac{1}{8x} + 2\left[\frac{a}{x} - \frac{a^2}{x}\right]\right) = \frac{1}{4x}(-2a^2 + 2a - 3/8)$$

After making the change of variables $dx = \sigma_X(x)dz = \sqrt{x}dz$ (thus $x = z^2/4$) we arrive at

$$J_X(z) = z^{-2}(-2a^2 + 2a - 3/8).$$

Similarly for the process Y_t

$$dy = \sigma_Y(y)dz = y^\theta dz \Rightarrow y = ((1-\theta)z)^{\frac{1}{1-\theta}}$$

and

$$J_Y(z) = \frac{1}{4}\left(\theta(\theta-1)y^{2\theta-2} - \frac{1}{2}\theta^2 y^{2\theta-2}\right) = \frac{1}{4} z^{-2} \frac{\frac{\theta^2}{2} - \theta}{(1-\theta)^2}.$$

One can check that when $\theta = 1 - \frac{1}{2(2a-1)}$ we have $J_Y(z) = J_X(z)$, thus processes Y_t should be related to the Bessel process by a stochastic transformation $\{\rho, h, Y\}$ with $\rho = 0$. We will construct this stochastic transformation explicitly later in section 3.4.2.

3.4 Examples

In this section we illustrate with some real examples the usefulness of stochastic transformations. We will review some well known examples (geometric Brownian motion, quadratic volatility family, CEV processes) and show how these processes can be obtained by a stochastic transformation, and in the case of Ornstein-Uhlenbeck, CIR and Jacobi processes we will construct new families of solvable driftless processes and study their boundary behavior.

3.4.1 Brownian Motion

Let $X_t = W_t$ be the Brownian motion process. Then the Markov generator is

$$\mathcal{L}^X = \frac{1}{2}\frac{d^2}{dx^2}.$$

Fix $\rho > 0$. Functions φ_ρ^+ and φ_ρ^- are given by:

$$\varphi_\rho^+(x) = e^{\sqrt{2\rho}x}, \qquad \varphi_\rho^-(x) = e^{-\sqrt{2\rho}x} \tag{3.4.1}$$

The next step is to fix any two positive c_1, c_2 and set $h = c_1\varphi_\rho^+ + c_2\varphi_\rho^-$. Note that $e^{-\rho t}h(X_t)$ is a martingale (a sum of two geometric brownian motions).

We consider separately two cases - (a) one of c_1, c_2 is zero (D_y is unbounded) and (b) both c_1, c_2 are positive.

Let's assume first that $c_1 = 0$ and $c_2 = 1$, thus

$$h(x) = \varphi_\rho^-(x) = e^{-\sqrt{2\rho}x}.$$

Thus the function $Y(x)$ is

$$Y(x) = \frac{c_3\varphi_\rho^+(x) + c_4\varphi_\rho^-(x)}{h(x)} = c_3 e^{2\sqrt{2\rho}x} + c_4,$$

and its inverse $x = X(y)$ is

$$X(y) = \frac{1}{2\sqrt{2\rho}} \log\left(\frac{|y - c_4|}{|c_3|}\right).$$

Using formula 3.3.2 we find the volatility function of the process Y_t:

$$\sigma_Y(y) = \sigma_X(X(y))\frac{1}{X'(y)} = C(y - y_1), \tag{3.4.2}$$

and $Y_t - y_1$ is the well known geometric brownian motion.

Note that this process is a martingale, it is transient and $\lim_{t\to\infty} Y_t = y_1$ a.s.

Now let's consider the general case: $c_1 > 0$, $c_2 > 0$. Assume that $c_4/c_2 > c_3/c_1$. Then

$$Y(x) = \frac{c_3\varphi_\rho^+(x) + c_4\varphi_\rho^-(x)}{c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x)} = \frac{c_3 e^{2\sqrt{2\rho}x} + c_4}{c_1 e^{2\sqrt{2\rho}x} + c_2},$$

and

$$X(y) = \frac{1}{2\sqrt{2\rho}}\left(\log\left(\frac{|c_2|}{|c_1|}\right) + \log\left(|c_4/c_2 - y|\right) - \log\left(|y - c_3/c_1|\right)\right).$$

Thus the derivative $X'(y)$ is

$$X'(y) = \frac{1}{C(c_4/c_2 - y)(y - c_3/c_1)},$$

and again using formula (3.3.2) we find the volatility function of the process Y_t

$$\sigma_Y(y) = C(c_4/c_2 - y)(y - c_3/c_1) = C(y_2 - y)(y - y_1), \quad y_2 > y_1. \tag{3.4.3}$$

and we obtain the quadratic volatility family. In this case the process Y_t is a uniformly integrable martingale. As $t \to \infty$ the process Y_t converges to the random variable Y_∞ with distribution supported on the boundaries y_1, y_2 and given by equation (3.3.7).

Remark 3.4.1. We have proved so far that starting from Brownian motion we can obtain the martingale processes with volatility function

$$\sigma_Y(y) = a_2 y^2 + a_1 y + a_0,$$

where the polynomial $a_2y^2 + a_1y + a_0$ is either linear or has two real zeros. However theorem 3.3.8 tells us that the process Y_t is related to W_t for all choices of coefficients a_i. Let's consider the example of the volatility function

$$\sigma_Y(y) = 1 + y^2 \tag{3.4.4}$$

to understand what happens in this case.

The process Y_t with volatility (3.4.4) is supported on the whole real line $\mathbb{R}$ and one can prove using Khasminskii's explosion test that this process explodes in finite time. We can obtain this process starting from Brownian motion by an analog of the stochastic transformation with $\rho < 0$. For example, let $\rho = -\frac{1}{2}$. Then solutions to equation

$$\mathcal{L}_X\varphi(x) = \frac{1}{2}\frac{d^2\varphi(x)}{dx^2} = \rho\varphi(x)$$

are given by functions $\sin(x)$ and $\cos(x)$. Let's choose

$$h(x) = \cos(x), \quad Y(x) = \frac{\sin(x)}{h(x)} = \tan(x).$$

Then $Y'(x) = \frac{1}{\cos^2(x)} = 1 + \tan^2(x) = 1 + y^2$, and we obtain the volatility function given in (3.4.4). Function $Y(x) = \tan(x)$ is infinite when $x = \pi/2 + k\pi$, $k \in \mathbb{Z}$, thus the process $Y_t = (Y(W_t), \mathrm{P}^h)$ explodes in finite time.

Thus we have proved the following

Lemma 3.4.2. *Starting with Brownian motion W_t we can obtain the class of quadratic volatility models:*

$$\mathfrak{M}(W_t) = \{Y_t : \sigma_Y(y) = a_2y^2 + a_1y + a_0\}. \tag{3.4.5}$$

3.4.2 Bessel processes

Let X_t be a Bessel process defined by generator

$$\mathcal{L}^X = a\frac{d}{dx} + \frac{1}{2}\sigma^2 x\frac{d^2}{dx^2}.$$

We assume that $\alpha = \frac{2a}{\sigma^2} - 1 > 0$ (thus the process never reaches the boundary point $x = 0$). Then we prove the following:

Lemma 3.4.3. *As a particular case one can cover the CEV (constant-elasticity-of-variance) models with volatility function $\sigma_Y(y) = c(y - y_0)^\theta$.*

Proof. Choose $\rho = 0$. Then the eigenvalue equation is

$$a\frac{d\varphi(x)}{dx} + \frac{1}{2}\sigma^2 x\frac{d^2\varphi(x)}{dx^2} = \rho\varphi(x) = 0.$$

The two linearly independent solutions are $\varphi_0^+(x) = 1$ and $\varphi_0^-(x) = x^{-\alpha}$. We put $c_2 = 0$, thus $h(x) = x^{-\alpha}$ and $Y(x)$ is

$$Y(x) = \frac{c_3x^{-\alpha} + c_4}{c_1x^{-\alpha}} = A + Bx^\alpha.$$

Thus we can express $x = X(y) = c_1(y - y_0)^{\frac{1}{\alpha}}$ and we find that

$$Y'(X(y)) = (X'(y))^{-1} = c_2(y - y_0)^{1-\frac{1}{\alpha}}.$$

Note that since $\alpha > 1$, the power $1 - \frac{1}{\alpha}$is positive. Using formula (3.3.2) we can compute the volatility

$$\sigma_Y(y) = \sigma\sqrt{X(y)}Y'(X(y)) = c(y - y_0)^{\theta},$$

where $\theta = 1 - \frac{1}{2}\frac{1}{\alpha}$. □

Remark 3.4.4. Note that we were able to compute the explicit form in these cases because that we could find the inverse function $x = X(y)$ explicitly. This is not the case for most applications, though, and we often have to use numerical inversion instead.

3.4.3 Ornstein-Uhlenbeck processes - OU family of martingales

Let X_t be the Ornstein-Uhlenbeck process

$$dX_t = (a - bX_t)dt + \sigma dWt,$$

discussed in detail in section 2.2.1.

Without loss of generality we can assume that $a = 0$ (otherwise we can consider the process $X_t - \frac{a}{b}$). Process X_t satisfies the following property, X_t has the same distribution as $-X_t$:

$$\mathrm{P}_x(X_t \in A) = \mathrm{P}_{-x}(-X_t \in A). \tag{3.4.6}$$

Functions $\varphi_\rho^-(x)$ and $\varphi_\rho^+(x)$ are solutions to the ODE:

$$\frac{1}{2}\sigma^2\frac{d^2\varphi(x)}{dx^2} - bx\frac{d\varphi(x)}{dx} = \rho\varphi(x). \tag{3.4.7}$$

One can check that the function $\varphi_\rho^-(x)$ is given by

$$\varphi_\rho^-(x) = \sqrt{\pi}\left(\frac{M(\frac{\rho}{2b}, \frac{1}{2}, \frac{b}{\sigma^2}x^2)}{\Gamma(\frac{1}{2} + \frac{\rho}{2b})} - 2\sqrt{\frac{b}{\sigma^2}}x\frac{M(\frac{\rho}{2b} + \frac{1}{2}, \frac{3}{2}, \frac{b}{\sigma^2}x^2)}{\Gamma(\frac{\rho}{2b})}\right) \tag{3.4.8}$$

and due to the symmetry of X_t (see (3.4.6)) we have $\varphi_\rho^+(x) = \varphi_\rho^-(-x)$.

Remark 3.4.5. To prove formula (3.4.8) one would start with the Kummer's equation (1.3.19) for $M(\frac{\rho}{2b}, \frac{1}{2}, z)$ and by the change of variables $z = \frac{b}{\sigma^2}x^2$ reduce this equation to the form (3.4.7).

We see that for $x > 0$ the function $\varphi_\rho^-(x)$ is just $U(\frac{\rho}{2b}, \frac{1}{2}, \frac{b}{\sigma^2}x^2)$, where U is the second solution to the Kummer's differential equation (1.3.22). As we see in the next section, functions M and U are related to the CIR process , since the square of Ornstein-Uhlenbeck $Y_t = X_t^2$ is a particular case of CIR process:

$$dY_t = (\sigma^2 - 2bY_t)dt + 2\sigma\sqrt{Y_t}dW_t.$$

The asymptotics of $\varphi_\rho^+(x)$ as $x \to \infty$ can be found using formula (1.3.24):

$$\varphi_\rho^+(x) \sim Cx^{\frac{\rho}{b}-1}e^{\frac{b}{\sigma^2}x^2}, \text{ as } x \to \infty; \quad \varphi_\rho^+(x) \sim Cx^{-\frac{\rho}{b}}, \text{ as } x \to -\infty \tag{3.4.9}$$

The asymptotics of $\varphi_\rho^-(x) = \varphi_\rho^+(-x)$ is obvious.

The boundary behavior of the process X^h and Y is given by the following lemma:

Lemma 3.4.6. *Let $h(x) = c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x)$. Then both boundaries are natural for the process X_t^h (the same for the process Y_t).*

Proof. Let's prove the result for the right boundary $D^2 = \infty$. The result for $D^1 = -\infty$ follows by symmetry.

Let $c_1 > 0$. From lemma 3.2.6 and equations (2.2.2) we find that the speed measure and the scale function of X^h have the following asymptotics as $x \to \infty$:

$$m_{X^h}(x) = h^2(x)m_X(x) \sim Cx^{2q}e^{\frac{b}{\sigma^2}x^2}, \quad s'_{X^h}(x) = h^{-2}(x)s'_X(x) \sim Cx^{-2q}e^{-\frac{b}{\sigma^2}x^2},$$

where $q = \frac{\rho}{b} - 1$. Now using lemma 1.1.5 one can check that $D^2 = \infty$ is a natural boundary. The case $c_1 = 0$ and $c_2 > 0$ can be analyzed similarly. □

Thus we see that the processes Y_t associated with the Ornstein-Uhlenbeck process behave similar to processes associated with Brownian motion (quadratic volatility family): these are conservative processes. The next two examples of processes illustrate different boundary behavior: the family associated with the CIR process has one killing (exit) boundary, while the family associated with the Jacobi process has both killing (exit) boundaries.

3.4.4 CIR processes - confluent hypergeometric family of driftless processes

Let X_t be the CIR process

$$dX_t = (a - bX_t)dt + \sigma\sqrt{X_t}dWt,$$

considered in section 2.2.3. We will use the notations from section 2.2.3:

$$\alpha = \frac{2a}{\sigma^2} - 1 \quad \text{and} \quad \theta = \frac{2b}{\sigma^2}.$$

Functions φ_ρ^+ and φ_ρ^- for the CIR process are solution to the ODE

$$\frac{1}{2}\sigma^2 x \frac{d^2\varphi(x)}{dx^2} - (a - bx)\frac{d\varphi(x)}{dx} = \rho\varphi(x). \tag{3.4.10}$$

Making affine change of variables $y = \theta x$ and dividing this equation by b, we reduce equation (3.4.10) to the Kummer differential equation (1.3.19), thus φ_ρ^+ and φ_ρ^- satisfy

$$\varphi_\rho^+(x) = M(\frac{\rho}{b}, \alpha + 1, \theta x), \tag{3.4.11}$$

$$\varphi_\rho^-(x) = U(\frac{\rho}{b}, \alpha + 1, \theta x). \tag{3.4.12}$$

Using formulas (1.3.22) and (1.3.24) we find the asymptotics of $\varphi_\rho^+(x)$ and $\varphi_\rho^-(x)$:

$$\varphi_\rho^+(x) \sim 1, \text{ as } x \to 0; \quad \varphi_\rho^+(x) \sim Ce^{\theta x} x^{\frac{\rho}{b} - \alpha - 1}, \text{ as } x \to \infty; \tag{3.4.13}$$

$$\varphi_\rho^-(x) \sim Cx^{-\alpha}, \text{ as } x \to 0; \quad \varphi_\rho^-(x) \sim Cx^{-\frac{\rho}{b}}, \text{ as } x \to \infty. \tag{3.4.14}$$

The following lemma describes the boundary behavior of the process X^h (and thus of the process $Y_t = Y(X^h)$).

Lemma 3.4.7. *Let $h(x) = c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x)$, where both c_i are positive. Then $D^2 = \infty$ is a natural boundary of the process X^h, while $D^1 = 0$ is a killing boundary if $\alpha \in (0, 1)$ and it is an exit boundary if $\alpha \geq 1$.*

Proof. Using formulas (3.4.13),(2.2.16) and lemma 3.2.6, we find that the asymptotics of the speed measure and scale function of X^h are

$$m_{X^h}(x) = h^2(x)m_X(x) \sim \begin{cases} Cx^{\frac{2\rho}{b} - \alpha - 2} e^{\theta x}, & \text{as } x \to \infty \\ Cx^{-\alpha}, & \text{as } x \to 0, \end{cases}$$

and

$$s'_{X^h}(x) = h^{-2}(x)s'_X(x) \sim \begin{cases} Cx^{-\frac{2\rho}{b} + \alpha - 1} e^{-\theta x}, & \text{as } x \to \infty \\ Cx^{\alpha - 1}, & \text{as } x \to 0. \end{cases}$$

One can check using Feller's theorem 1.1.5, that for $\alpha \in (0, 1)$ we have a regular boundary, but since the h-transform introduces nonzero killing measure at the boundaries given by equation (3.2.9), we have a killing boundary. If $\alpha \geq 1$ we have an exit boundary. □

Note that since the left boundary $D^1 = 0$ is either a killing or an exit boundary, the process X^h is not a conservative process. The same is true for $Y_t = Y(X_t^h)$.

Theorem 3.4.8. *The family of driftless processes Y_t related to a CIR process by a stochastic transformation is characterized by their volatility functions as follows:*

$$\sigma_Y(Y(x)) = C\sqrt{x}\frac{x^{-\alpha-1}e^{\theta x}}{(c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x))^2}, \tag{3.4.15}$$
$$Y(x) = \frac{c_3\varphi_\rho^+(x) + c_4\varphi_\rho^-(x)}{c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x)}.$$

Definition 3.4.9. We will call this family of driftless processes *the confluent hypergeometric family* and denote it by

$$\mathfrak{M}(\text{CIR}) = \{Y_t : Y \sim \text{CIR process}\}.$$

3.4.5 Jacobi processes - hypergeometric family of driftless processes

Let X_t be the Jacobi process

$$dX_t = (a - bX_t)dt + \sigma\sqrt{X_t(A - X_t)}dW_t,$$

described in section 2.2.5. Recall the notations from section 2.2.5. We defined parameters

$$\alpha = \frac{2b}{\sigma^2} - \frac{2a}{\sigma^2 A} - 1 \text{ and } \beta = \frac{2a}{\sigma^2 A} - 1.$$

We assume that $\alpha > 0$ and $\beta > 0$, thus both boundaries are inaccessible (see section 2.2.5).

Functions φ_ρ^+ and φ_ρ^- for the Jacobi process are solutions to the ODE

$$\frac{1}{2}\sigma^2 x(A - x)\frac{d^2\varphi(x)}{dx^2} - (a - bx)\frac{d\varphi(x)}{dx} = \rho\varphi(x).$$

By the affine change of variables $x = Ay$ this equation is reduced to the hypergeometric differential equation (1.3.9), thus using equations (1.3.16) we find that functions φ_ρ^+ and φ_ρ^- for the Jacobi process are given by

$$\varphi_\rho^+(x) = {}_2F_1(\alpha_1, \alpha_2; \beta_1; x/A) \tag{3.4.16}$$

$$\varphi_\rho^-(x) = {}_2F_1(\alpha_1, \alpha_2; \alpha_1 + \alpha_2 + 1 - \beta_1; 1 - x/A) \tag{3.4.17}$$

where the parameters satisfy the following system of equations:

$$\begin{cases} \alpha_1 + \alpha_2 + 1 = \frac{2b}{\sigma^2} \\ \alpha_1 \alpha_2 = \frac{2\rho}{\sigma^2} \\ \beta_1 = \frac{2a}{A\sigma^2}. \end{cases} \tag{3.4.18}$$

The asymptotics of $\varphi_\rho^+(x)$ and $\varphi_\rho^-(x)$ are

$$\varphi_\rho^+(x) \sim 1, \text{ as } x \to 0, \quad \varphi_\rho^+(x) \sim C(A-x)^{-\alpha}, \text{ as } x \to A, \tag{3.4.19}$$

$$\varphi_\rho^-(x) \sim Cx^{-\beta}, \text{ as } x \to 0, \quad \varphi_\rho^-(x) \sim 1, \text{ as } x \to A. \tag{3.4.20}$$

Let $h(x) = c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x)$, where both c_i are positive. The boundary behavior of the h-transformed process X_t^h is similar to the case of CIR process at $x = 0$:

Lemma 3.4.10. *$D^1 = 0$ is a killing boundary for the process X_t^h if $\beta \in (0,1)$ and it is an exit boundary if $\beta \geq 1$. The same is true for $D^2 = A$ by changing $\beta \mapsto \alpha$.*

We see that in the case of the Jacobi process both boundaries are either killing or exit boundaries for X_t^h, thus X_t^h and Y_t are not conservative processes.

Theorem 3.4.11. *The family of driftless processes Y_t related to a Jacobi process by a stochastic transformation is characterized by their volatility functions as follows:*

$$\sigma_Y(Y(x)) = C\sqrt{x(A-x)}\frac{x^{-\alpha-1}(A-x)^{-\beta-1}}{(c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x))^2}, \tag{3.4.21}$$
$$Y(x) = \frac{c_3\varphi_\rho^+(x) + c_4\varphi_\rho^-(x)}{c_1\varphi_\rho^+(x) + c_2\varphi_\rho^-(x)}.$$

Definition 3.4.12. We call this family of driftless processes the *the hypergeometric family* and denote it by

$$\mathfrak{M}(\text{Jacobi}) = \{Y_t : Y \sim \text{Jacobi process}\}.$$

3.5 Discrete phase space - birth and death processes

In this section we construct discrete approximations to the processes discussed above.

For $\delta > 0$, let ∇_δ^+ be the forward finite difference operator

$$\nabla_\delta^+ f(x) = \frac{f(x+\delta) - f(x)}{\delta} \tag{3.5.1}$$

and let Δ_δ be the discretized Laplace operator

$$\Delta_\delta f(x) = \frac{f(x+\delta) - 2f(x) + f(x-\delta)}{\delta^2} \tag{3.5.2}$$

The finite difference equivalents of the generator $\mathcal{L}_X$ are Markov generators of discrete state Markov processes and have the form

$$\mathcal{L}_X^\delta = b(x)\nabla_\delta^+ + \frac{\sigma(x)^2}{2}\Delta_\delta, \tag{3.5.3}$$

where $b(x)$ and $\sigma(x)$ are functions defined on the lattice

$$\Lambda = \Lambda(\delta, D) := [D^1, D^2] \cap \{D^1 + \delta\mathbb{Z}\} = \{D^1, D^1 + \delta, D^1 + 2\delta, \dots\}. \tag{3.5.4}$$

If D^1 (and/or D^2) is finite, define the interior of $\Lambda(\delta, D)$ as the set $\Lambda(\delta, D)$ without D^1 (and/or D^2):

$$\text{int}\Lambda := \Lambda(\delta, D) \setminus \{D^1, D^2\}. \tag{3.5.5}$$

Theorem 3.5.1. *Fix a positive δ and let $D_x = \Lambda(\delta, D)$ be a domain on the lattice. Let X_t^δ be a birth and death process under the measure P taking values on the domain $\Lambda(h, D)$ and admitting a Markov generator $\mathcal{L}_X^\delta$ of the form in (3.5.3). Select $\rho > 0$ and let $\varphi_1(x)$, $\varphi_2(x)$ be the two linearly independent solutions to the finite difference equation*

$$\mathcal{L}_X^\delta \varphi(x) = \rho\varphi(x), \quad x \in \mathit{int}\Lambda. \tag{3.5.6}$$

Define the function

$$h(x) = c_1\varphi_1(x) + c_2\varphi_2(x), \tag{3.5.7}$$

where c_1 and c_2 are two positive constants such that $h(X_0^\delta) = 1$. Then we have the following:

(i) There exists a measure P^h absolutely continuous with respect to P defined by

$$\frac{d\mathrm{P}_t^h}{d\mathrm{P}_t} = e^{-\rho t} h(X_t^\delta). \tag{3.5.8}$$

(ii) Consider the function

$$Y(x) = \frac{c_3\varphi_1(x) + c_4\varphi_2(x)}{c_1\varphi_1(x) + c_2\varphi_2(x)}. \tag{3.5.9}$$

Assume that for all $x \in D_x$

$$\frac{2\delta b(x)}{\sigma^2(x)} \le 1. \tag{3.5.10}$$

Then function $Y(x)$ is invertible for all choices of c_i such that $c_1c_4 - c_2c_3 \neq 0$. Define the process $Y_t^\delta = (Y(X_t^\delta), \mathrm{P}^h)$. If T is the stopping time at which X_t^δ hits the boundary of the domain Λ, then the process Y^T is a bounded martingale.

Proof. The arguments used in section 3.2.3 extend to the discrete case. We only show that function $Y(x)$ is invertible. Let $g(x) = c_3\varphi_1(x) + c_4\varphi_2(x)$, thus $Y(x) = g(x)/h(x)$.

$$Y(x+\delta) - Y(x) = \frac{g(x+\delta)h(x) - h(x+\delta)g(x)}{h(x)h(x+\delta)} = \frac{W_{g,h}(x)}{h(x)h(x+\delta)},$$

where $W(x) = W_{g,h}(x) := g(x+\delta)h(x) - h(x+\delta)g(x)$ is a discrete analog of the Wronskian. Since functions g and h are solutions to the finite difference equation

$$\mathcal{L}_X^\delta\varphi(x) = b(x)\nabla_\delta^+\varphi(x) + \frac{\sigma(x)^2}{2}\Delta_\delta\varphi(x) = \rho\varphi(x),$$

one can check by direct computation that the Wronskian satisfies the following difference equation:

$$\nabla_\delta^+ W(x) = \frac{W(x+\delta) - W(x)}{\delta} = -\frac{2b(x)}{\sigma^2(x)}W(x).$$

Thus $W(x+\delta) = W(x)(1-\delta\frac{2b(x)}{\sigma^2(x)})$, so under condition (3.5.10) either $W(x) \equiv 0$ or it is never zero. Since g and h are independent solutions, the Wronskian is never zero, thus $Y(x+\delta) - Y(x)$ never changes sign and the function $Y(x)$ is monotonous, thus invertible. □

The following lemma (an analog of lemma 3.2.6) summarizes the important properties of X^δ:

Lemma 3.5.2. *The process $X_t^{\delta,h} = (X_t^\delta, \mathrm{P}^h)$ has the generator*

$$\mathcal{L}_{X^h}^\delta = \frac{1}{h}\mathcal{L}_X^\delta h - \rho. \tag{3.5.11}$$

The transition probability of Y_t^δ and $X_t^{\delta,h}$ are given by

$$p_{Y^\delta}(t, y_0, y_1) = p_{X^{\delta,h}}(t, x_0, x_1) = e^{-\rho t}\frac{h(x_1)}{h(x_0)}p_{X^\delta}(t, x_0, x_1), \tag{3.5.12}$$

where $y_i = Y(x_i)$.

Chapter 4

Classification of solvable driftless diffusions

In this chapter we raise the idea of stochastic transformations to a new level. Instead of reducing Markov process to other solvable processes, we use different tools (canonical forms of ODEs, Liouville transformations and Bose invariants) to map "eigenfunction" equation for the generator of Markov driftless process into the (confluent) hypergeometric equation. This lets us compute the Green function in terms of (confluent) hypergeometric functions, thus the probability density can be computed as an inverse Laplace transform of the Green function (or as an integral over hypergeometric functions).

In the first section we define the Liouville transformation as a way to map an equation in canonical (selfadjoint) form into another canonical form. We discuss several properties of the Liouville transformation with the emphasis on similarities and differences between the former and stochastic transformations.

In the second section we state and prove two classification theorems – the first theorem gives an explicit formula for the volatility function of the driftless process solvable in terms of (confluent) hypergeometric functions. This formula is a generalization of formulas (3.4.15) and (3.4.21). The second classification theorem shows that these processes can be constructed by stochastic transformations starting from some simpler solvable processes (which is again just a generalization of the construction given in sections 3.4.4 and 3.4.5 for the CIR and Jacobi processes.

4.1 Introduction

In the introduction to chapter 3 we discuss the transformations that preserve the solvability of the Markov process. In this chapter we work only with the second order differential operators and solutions to the "eigenfunction" equation. Let's ask a question: how can one transform operator $\mathcal{L}$ and solutions to the eigenfunction equation?

Let $\mathcal{L}$ be the second order differential operator

$$\mathcal{L} = a(y)\frac{\partial^2}{\partial y^2} + b(y)\frac{\partial}{\partial y}. \tag{4.1.1}$$

If $a(y)$ is positive on some interval D (and $a(y)$ and $b(y)$ satisfy some other conditions), operator $\mathcal{L}$ can be considered as a generator of a diffusion process, thus we will call $a(y)$ the *volatility coefficient* and $b(y)$ the *drift coefficient* of operator $\mathcal{L}$.

Let's consider the following three types of transformations

(i) Change of variables: $x = x(y)$, $f(y) \mapsto f(y(x))$ and

$$\mathcal{L} \mapsto T_{y\to x}\mathcal{L} \;= a(y)(x'(y))^2\frac{\partial^2}{\partial x^2} + (a(y)x''(y) + b(y)x'(y))\frac{\partial}{\partial x} = \tag{4.1.2}$$

$$a(y)(x'(y))^2\frac{\partial^2}{\partial x^2} + (\mathcal{L}x)(y)\frac{\partial}{\partial x}, \tag{4.1.3}$$

where $y = y(x)$.

(ii) Gauge transformation: $f(y) \mapsto h(y)f(y)$ and

$$\mathcal{L} \mapsto T_h\mathcal{L} = \tfrac{1}{h}\mathcal{L}h \;= a(y)\frac{\partial^2}{\partial y^2} + \left(b(y) + 2a(y)\frac{h'(y)}{h(y)}\right)\frac{\partial}{\partial y} + \frac{1}{h}\left(a(y)h''(y) + b(y)h'(y)\right) =$$

$$= a(y)\frac{\partial^2}{\partial y^2} + \left(b(y) + 2a(y)\frac{h'(y)}{h(y)}\right)\frac{\partial}{\partial y} + \frac{1}{h(y)}(\mathcal{L}h)(y). \tag{4.1.4}$$

Notice that gauge transformation actually consists of two transformations: right multiplication of $\mathcal{L}$ by h and left multiplication by $1/h$.

(iii) Left multiplication by a function $\gamma^2(x)$ (which does not change $f(x)$):

$$\mathcal{L} \mapsto T_{\gamma^2}\mathcal{L} = \gamma^2(y)a(y)\frac{\partial^2}{\partial y^2} + \gamma^2(y)b(y)\frac{\partial}{\partial y}. \tag{4.1.5}$$

In the case when $a(y) = \frac{1}{2}\sigma^2(y)$ and $\mathcal{L}$ is a generator of a Markov diffusion (Y_t, P), we can give a probabilistic interpretation to the first and second types of transformations described above: $T_{y\to x}$ is just the usual change of variables formula for the stochastic process Y_t, which describes the dynamics of the process $X_t = X(Y_t)$, thus it is just an analog of the Ito formula written in the language of ODEs. The Gauge transformation has a probabilistic meaning if $h(Y_t)$ can be considered as a measure change density (thus $h(Y_t)$ is a local martingale and $\mathcal{L}h = 0$), or when h is a ρ-excessive function ($\mathcal{L}h = \rho h$) – then the gauge transformation $\frac{1}{h}\mathcal{L}h - \rho$ is the Doob's h-transform discussed in section 3.2.1. The transformed generator $\mathcal{L}$ in this case describes the dynamics of the process Y_t under the new measure Q, defined by $d\mathrm{Q}_t = h(Y_t)d\mathrm{P}_t$. Note that both of these transformations preserve the form of backward Kolmogorov equation:

$$\frac{\partial}{\partial t}f(t,y) = \mathcal{L}f(t,y).$$

The last transformation T_{γ^2} does not preserve the form of backward Kolmogorov equation, thus in general it has no immediate probabilistic meaning, except when $\gamma^2(x) = c$ is constant – then T_{γ^2} is equivalent to scaling of time $t \mapsto \frac{1}{c}t'$.

As we have seen in chapter 3, the CIR and Jacobi families of processes are solvable because they can be reduced to some simple solvable process. In other words, for these processes, the eigenfunction equation

$$\mathcal{L}_Y f(y) = \frac{1}{2}\sigma_Y^2(y)\frac{\partial^2 f(y)}{\partial y^2} = \rho f(y) \tag{4.1.6}$$

can be reduced by a gauge transformation (change of measure) and a change of variables to a hypergeometric or a confluent hypergeometric equation, thus giving us eigenfunctions $\psi_n(x)$, generalized eigenfunctions φ_λ^+, φ_λ^-, and a possibility to compute the transitional probability density as

$$p_Y(t,y_0,y_1) = \frac{e^{-\rho t}}{h(x_0)h(x_1)}p_X(t,x_0,x_1) = \frac{1}{h(x_0)h(x_1)}\sum_{n=0}^{\infty} e^{-(\rho-\lambda_n)t}\psi_n(x_0)\psi_n(x_1). \tag{4.1.7}$$

It is known that OU, CIR and Jacobi processes are the only diffusions associated with a system of orthogonal polynomials (see [20]), thus corresponding families of driftless processes are the only ones that can have a probability density of the form (4.1.7) where the orthogonal basis $\{\psi_n\}_{n\geq 0}$ is given by orthogonal polynomials. However we might hope to find new families of solvable processes if we generalize the definition of solvability.

Note that the fact that we can compute solutions of equation (4.1.6) gives us a ready expression for the Green function through the formula (1.1.22):

$$G_Y(\lambda,y_0,y_1) = \begin{cases} w_\lambda^{-1}\varphi_\lambda^+(y_0)\varphi_\lambda^-(y_1), & y_0 \leq y_1 \\ w_\lambda^{-1}\varphi_\lambda^+(y_1)\varphi_\lambda^-(y_0), & y_1 \leq y_0. \end{cases} \tag{4.1.8}$$

Since the Green function is a Laplace transform of $p_Y(t,y_0,y_1)$ one could hope to find the probability kernel through the inverse Laplace transform of $G_Y(\lambda,y_0,y_1)$. Thus in this chapter we will use the following definition of solvability:

Definition 4.1.1. The one dimensional diffusion process Y_t on the interval D_y is called *solvable*, if its Green function can be computed in terms of (scaled confluent) hypergeometric functions.

In other words, the process Y_t is solvable if there exist a λ-independent change of variables $y = y(z)$ and a (possibly λ-dependent) function $h(z, \lambda)$, such that all solutions to the equation

$$\mathcal{L}_Y f(y) = \lambda f(y) \tag{4.1.9}$$

are of the form $h(z(y), \lambda)F(z(y))$, where F is either a hypergeometric function ${}_2F_1(a, b; c; z)$, or a scaled confluent hypergeometric function $M(a, b, wz)$ (with parameters depending on λ).

Remark 4.1.2. Later we will see that the requirement that the change of variables is independent of λ is necessary, because otherwise every diffusion process is solvable (see remark (4.3.5)).

4.2 Liouville transformations and Bose invariants

Consider the linear second order differential operator

$$\mathcal{L}_y = a(y)\frac{\partial^2}{\partial y^2} + b(y)\frac{\partial}{\partial y}. \tag{4.2.1}$$

Making the gauge transformation with gauge factor h

$$h(y) = \exp\left(-\int^y \frac{b(u)}{2a(u)}du\right) = \sqrt{W(y)},$$

we can remove the “drift” coefficient and thereby arrive at the operator

$$\mathcal{L}_y \mapsto \frac{1}{h}\mathcal{L}_y h = a(y)\frac{\partial^2}{\partial y^2} + a(y)I(y).$$

Multiplying this operator by $a(y)^{-1}$ we arrive at symmetric operator

$$\frac{1}{h}\mathcal{L}_y h \mapsto a^{-1}\frac{1}{h}\mathcal{L}_y h = \mathcal{L}^c = \frac{\partial^2}{\partial y^2} + I(y), \tag{4.2.2}$$

where the potential term is given by:

$$I(y) = \left(\frac{h'(y)}{h(y)}\right)' - \left(\frac{h'(y)}{h(y)}\right)^2 = \frac{2b(y)a(y)' - 2a(y)b(y)' - b(y)^2}{4a(y)^2}. \tag{4.2.3}$$

Definition 4.2.1. $\mathcal{L}^c$ is called the *canonical form* of the operator $\mathcal{L}_y$.

The form of operator $\mathcal{L}^c$ is clearly invariant with respect to any of the three types of transformations described above. Moreover, the following lemma is true:

Lemma 4.2.2. *The canonical form of the operator $\mathcal{L}^c$ given by (4.2.2) is invariant under any two transformations of $\{T_{y\to z}, T_h, T_{\gamma^2}\}$.*

Proof. This lemma is proved by checking that every combination of two transformations cannot change the canonical form of the operator (4.2.2). For example, suppose we are free to use $T_{y\to z}$ and T_h, but not T_{γ^2}. Using formula (4.1.4) we find that by applying T_h we have nonzero drift given by $2h'/h$. We can't remove this drift by some change of variables $T_{y\to x}$, since by formula (4.1.2) it will add a nontrivial volatility term $x'(y)^2$, which in turn can not be removed by any gauge transformation T_h. As another example let's assume that we can use $T_{y\to x}$ and T_{γ^2}, but not T_h. By formula (4.1.5) we see that T_{γ^2} adds nontrivial volatility, which can be removed by $T_{y\to x}$, but formula (4.1.2) tells us that this in turn will add nonzero drift $\gamma^2(x)(\mathcal{L}x)(y)$, which can't be removed by any T_{γ^2}. The last combination T_{γ^2} and T_h can be checked in exactly the same way. □

Definition 4.2.3. Function $I(y)$ is called the *Bose invariant* of operator (4.2.1).

Remark 4.2.4. Note that to bring $\mathcal{L}_y$ to canonical form $\mathcal{L}^c$ we used two types of transformations: a gauge transformation and a change of variables. But by using other choices of two transformations we could bring $\mathcal{L}_y$ to a different canonical form. Thus when we talk about canonical form we always specify with respect to which two types of transformations this form is invariant (in this chapter we use only two types of canonical forms: one described above, and the second obtained by a gauge transformation and a change of variables only).

As we proved, function $I(y)$ is invariant with respect to any two transformations of $\{T_{y\to x}, T_h, T_{\gamma^2}\}$. However it is possible to change the potential $I(y)$ by applying all three types of transformations. The idea is to apply first a change of variables transformation, then remove the "drift" by a gauge transformation, after which we divide by the volatility coefficient to obtain a new canonical form. The details are:

(i) A change of the independent variable $y = y(x)$ (see equation (4.1.2)) changes the operator (4.2.2) into

$$\mathcal{L}_x = (y'(x))^{-2}\frac{\partial^2}{\partial x^2} - \frac{y''(x)}{(y'(x))^3}\frac{\partial}{\partial x} + I(y(x)).$$

(ii) Multiplying the operator $\mathcal{L}_x$ by $\gamma^2(x) = (y'(x))^2$ we arrive at

$$(y'(x))^2\mathcal{L}_x = \frac{\partial^2}{\partial x^2} - \frac{y''(x)}{y'(x)}\frac{\partial}{\partial x} + (y'(x))^2 I(y(x)),$$

(iii) Applying the gauge transformation with gauge factor $h = \sqrt{y'(x)}$ (see equation (4.1.4)) brings us to the operator in the following canonical form

$$\mathcal{L}_x^c = \frac{\partial^2}{\partial x^2} + J(x),$$

where the potential term is transformed as

$$J(x) = \frac{1}{2}\{y, x\} + (y'(x))^2 I(y(x)), \tag{4.2.4}$$

and $\{y, x\}$ is the *Schwarzian derivative of y with respect to x*:

$$\{y, x\} = \left(\frac{y''(x)}{y'(x)}\right)' - \frac{1}{2}\left(\frac{y''(x)}{y'(x)}\right)^2.$$

The above transformation changes the canonical form of the operator by applying all three types of transformations. It is called a *Liouville transformation*.

Note that the order of the different steps in Liouville transformation does not matter, since all the three transformations $\{T_{y\to x}, T_h, T_{\gamma^2}\}$ commute. The other important idea is that we are free to choose the first transformation, but the other two are uniquely defined by the first one (see lemma (4.2.2)). We will use this fact in the proof of the main theorem in the next section.

The following lemma is an immediate application of the above theory, which gives an answer to the following question: given a diffusion X_t with Markov generator $\mathcal{L}_X$ and a driftless process Y_t with generator $\mathcal{L}_Y$, under which conditions on generators of X and Y can we be sure that Y is related to X by a stochastic transformation?

Lemma 4.2.5. *Let X_t and Y_t be diffusion processes with generators*

$$\mathcal{L}_X = b_X(x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma_X^2(x)\frac{\partial^2}{\partial x^2}, \quad \mathcal{L}_Y = \frac{1}{2}\sigma_Y^2(y)\frac{\partial^2}{\partial y^2}. \tag{4.2.5}$$

Then the Bose invariants $J_X(z)$ and $J_Y(z)$ for the differential equations $\mathcal{L}_X f = \lambda f$ and $\mathcal{L}_Y f = \lambda f$ obtained by a change of variables and a gauge transformation can be computed as follows:

$$J_X(z(x)) = \frac{1}{4}\left(\sigma_X(x)\sigma_X''(x) - \frac{1}{2}(\sigma_X'(x))^2 + 2\left[\frac{2b_X(x)\sigma_X'(x)}{\sigma_X(x)} - b_X'(x) - \frac{b_X^2(x)}{\sigma_X^2(x)}\right]\right), \tag{4.2.6}$$

$$z'(x) = \frac{1}{\sigma_X(x)},$$

and the corresponding expression for $J_Y(z)$

$$J_Y(z(y)) = \frac{1}{4}\left(\sigma_Y(y)\sigma_Y''(y) - \frac{1}{2}(\sigma_Y'(y))^2\right), \qquad (4.2.7)$$
$$z'(y) = \frac{1}{\sigma_Y(y)}.$$

Functions $J_X(z)$ *and* $J_Y(z)$ *are invariant under change of variables and gauge transformations, thus process* Y_t *is related to* X_t *by some stochastic transformation* $\{\rho, h, Y\}$ *if and only if*

$$J_Y(z) \equiv J_X(z) - \rho. \qquad (4.2.8)$$

Proof. Applying the gauge transformation with gauge factor $h_1(x) = \sqrt{W(x)}$ changes $\mathcal{L}_X$ as follows:

$$\mathcal{L}_X \mapsto \frac{1}{h_1}\mathcal{L}_X h_1 = \frac{1}{2}\sigma_X^2(x)\frac{\partial^2}{\partial x^2} + I_1(x), \qquad (4.2.9)$$

where by formula (4.2.3) the potential term is equal

$$I_1(x) = \frac{1}{2}\left[\frac{2b_X(x)\sigma_X'(x)}{\sigma_X(x)} - b_X'(x) - \frac{b_X^2(x)}{\sigma_X^2(x)}\right]. \qquad (4.2.10)$$

Changing variables $x'(z) = \frac{1}{\sqrt{2}}\sigma_X(x)$ we arrive at

$$\mathcal{L}_z = \frac{\partial^2}{\partial z^2} - \frac{x''(z)}{x'(z)}\frac{\partial}{\partial z} + I_1(x(z)), \qquad (4.2.11)$$

and at last making the gauge transformation with gauge factor $h_2(z) = \sqrt{x'(z)}$ we arrive at the canonical form

$$\mathcal{L}^c = \frac{1}{h_2}\mathcal{L}_z h_2 = \frac{\partial^2}{\partial z^2} + J_X(z) \qquad (4.2.12)$$

where the Bose invariant (potential term) is

$$J_X(z) = \frac{1}{2}\{x, z\} + I_1(x). \qquad (4.2.13)$$

By direct computations we check that $\{x, z\} = \frac{1}{2}\sigma_X''(x)\sigma_X(x) - \frac{1}{4}\sigma_X^2(x)$, which gives us formula (4.2.6).

Now we need to prove that $J_X(z)$ is invariant under stochastic transformations up to an additive constant. A change of variables does not change the form of $J_X(z)$. A change of measure (ρ-excessive transform) changes the generator of X

$$\mathcal{L}_X \mapsto \frac{1}{h}\mathcal{L}_X h - \rho.$$

Since $J_X(z)$ is invariant under gauge transformations $\mathcal{L}_X \mapsto \frac{1}{h}\mathcal{L}_X h$ we see that $J_Y(z)$ differs from $J_X(z)$ by a constant $-\rho$, which ends the proof. □

The following diagram illustrates the meaning of a Liouville transformation (in particular the T_{γ^2} transformation). Assume that we start with a diffusion process X_1. By stochastic transformations we can construct the family $\mathfrak{M}(X_1)$ of driftless processes. By a gauge transformation and by a change of variables we map the generator of X_1 into the corresponding canonical form (with the Bose invariant given by (4.2.6)). By applying a Liouville transformation we change the potential (Bose invariant) of the canonical form and then by a change of variables and a gauge transformation we can map it back into the generator of some diffusion process X_2. Then as before we construct a family of driftless processes $\mathfrak{M}(X_2)$. Note that $\mathfrak{M}(X_2)$ and $\mathfrak{M}(X_1)$ are not related by a stochastic canonical transformation, since otherwise they would have the same Bose invariants. Then the process can be continued. Thus we have a family of Bose invariants related by Liouville transformations, each of these invariants (potentials in canonical form) gives rise to a new family of driftless processes, not related by a stochastic transformation to the previous families.

The following diagram summarizes the above ideas:

$\mathfrak{M}(X_1)$		$\mathfrak{M}(X_2)$		$\mathfrak{M}(X_3)$
$\wr$		$\wr$		$\wr$
stochastic transformations		stochastic transformations		stochastic transformations
$\wr$		$\wr$		$\wr$
X_1	$\not\sim$	X_2	$\not\sim$	X_3
$\Updownarrow$		$\Updownarrow$		$\Updownarrow$
$J_{X_1}(z)$	$\Longleftrightarrow$	$J_{X_2}(z)$	$\Longleftrightarrow$	$J_{X_3}(z)$
	Liouville transformation		Liouville transformation	

4.3 Classification: Main theorems

Theorem 4.3.1. First classification theorem. *A driftless process Y_t is solvable in the sense of definition 4.1.1 if and only if its volatility function is of the following form:*

$$\sigma_Y(y) = \sigma_Y(Y(x)) = C\sqrt{A(x)}\frac{W(x)}{(c_1F_1(x)+c_2F_2(x))^2}\sqrt{\frac{A(x)}{R(x)}}, \tag{4.3.1}$$

where the change of variables is given by

$$y = Y(x) = \frac{c_3 F_1(x) + c_4 F_2(x)}{c_1 F_1(x) + c_2 F_2(x)}, \quad c_1 c_4 - c_3 c_2 \neq 0. \tag{4.3.2}$$

In the case $A(x) = x$*:*

(i) $R(x) \in P_2$*, such that* $R(x) \neq 0$ *in* $(0, \infty)$

(ii) F_1 *and* F_2 *are functions* $M(a, b, wx)$ *and* $U(a, b, wx)$

(iii) $W(x)$ *is a Wronskian of the scaled Kummer differential equation (equal to* $s'(x)$ *in (2.2.16)).*

and in the case $A(x) = x(1 - x)$

(i) $R(x) \in P_2$*, such that* $R(x) \neq 0$ *in* $(0, 1)$

(ii) F_1 *and* F_2 *are two linearly independent solutions to the hypergeometric equation given by (1.3.11).*

(iii) $W(x)$ *is a Wronskian of the hypergeometric differential equation (equal to* $s'(x)$ *in (2.2.29)).*

Before we prove this theorem we need to establish some auxiliary results.

Lemma 4.3.2. *The Bose invariant for the hypergeometric equation*

$$x(1-x)\frac{\partial^2 f(x)}{\partial x^2} + (\gamma - (1 + \alpha + \beta)x)\frac{\partial f(x)}{\partial x} - \alpha\beta f(x) = 0.$$

is given by

$$I_{hyp}(x) = \frac{Q(x)}{4x^2(1-x)^2}, \tag{4.3.3}$$

where

$$Q(x) = (1 - (\alpha - \beta)^2)x^2 + (2\gamma(\alpha + \beta - 1) - 4\alpha\beta)x + \gamma(2 - \gamma). \tag{4.3.4}$$

The Bose invariant for the scaled confluent hypergeometric equation

$$x\frac{\partial^2 f(x)}{\partial x^2} + (a - wx)\frac{\partial f(x)}{\partial x} - bwf(x) = 0$$

is given by

$$I_{confl}(x) = \frac{Q(x)}{4x^2}, \tag{4.3.5}$$

where

$$Q(x) = -w^2x^2 + 2w(a-2b)x + a(2-a). \tag{4.3.6}$$

In both cases by varying parameters we can obtain any second order polynomial Q.

Proof. For the second order differential equation

$$a(x)\frac{\partial^2 f(x)}{\partial x^2} + b(x)\frac{\partial f(x)}{\partial x} + c(x)f(x) = 0 \tag{4.3.7}$$

the Bose invariant is given by formula (4.2.3) and is equal to

$$\begin{aligned} I(x) &= \frac{2b(x)a(x)' - 2a(x)b(x)' - b(x)^2}{4a(x)^2} + \frac{c(x)}{a(x)} \\ &= \frac{2b(x)a(x)' - 2a(x)b(x)' - b(x)^2 + 4a(x)c(x)}{4a(x)^2}. \end{aligned}$$

□

Corollary 4.3.3. *Let $T(x) \in P_2$ be an arbitrary second order polynomial and $A(x) \in \{x, x(1-x)\}$. The solutions to equation*

$$\frac{\partial^2 f(x)}{\partial x^2} + \frac{T(x)}{A^2(x)}f(x) = 0$$

can be obtained in the form $F(x)/\sqrt{W(x)}$, where $W(x)$ is the Wronskian given by $\exp(-\int^x \frac{b(y)}{a(y)}dy)$ and function $F(x)$ is a solution to the hypergeometric equation in the case $A(x) = x(1-x)$ or to the scaled confluent hypergeometric equation in the case $A(x) = x$.

Theorem 4.3.4. (Schwarz) *The general solution to the equation*

$$\frac{1}{2}\{y, x\} = J(x) \tag{4.3.8}$$

has the form

$$y(x) = \frac{F_2(x)}{F_1(x)}, \tag{4.3.9}$$

where F_1 and F_2 are arbitrary linearly independent solutions of equation

$$\frac{\partial^2 F(x)}{\partial x^2} + J(x)F(x) = 0. \tag{4.3.10}$$

Proof. Let's introduce the new variable $F(x) = \frac{1}{\sqrt{y'(x)}}$. Then we have

$$y'(x) = \frac{1}{(F(x))^2} \quad \Rightarrow \log y'(x) = -2\log(F(x)) \quad \Rightarrow \frac{y''(x)}{y'(x)} = -2\frac{F'(x)}{F(x)}.$$

Thus the Schwarzian derivative $\{y, z\}$ is equal to:

$$\{y, z\} = \left(\frac{y''(x)}{y'(x)}\right)' - \frac{1}{2}\left(\frac{y''(x)}{y'(x)}\right)^2 = -2\frac{F''(x)}{F(x)}$$

and equation (4.3.8) is transformed into equation

$$F''(x) + J(x)F(x) = 0$$

for function $F(x)$, thus

$$y'(x) = \frac{1}{(F_1(x))^2},$$

where $F_1(x)$ is an arbitrary solution of $F''(x) + J(x)F(x) = 0$.

Note that if F_1 and F_2 are solutions to second order linear ODE $aF'' + bF' + cF = 0$, then

$$\frac{d}{dx}\left(\frac{F_2(x)}{F_1(x)}\right) = \frac{F_2'(x)F_1(x) - F_2(x)F_1'(x)}{F_2^2(x)} = \frac{W_{F_2,F_1}(x)}{F_1^2(x)},$$

and since the Wronskian of equation (4.3.10) is constant we can obtain the above expression for $y(x)$. □

Remark 4.3.5. Note that theorem 4.3.4 and equation (4.2.4) tells us that for any potential there exists a change of variables $z(x)$, which maps equation $F'' = 0$ into equation $F''(x) + J(x)F(x) = 0$, thus any two canonical forms can be related by some Liouville transformation. That is why in the definition (4.1.1) we require the change of variables $y(x)$ to be independent of λ.

Proof of the first classification theorem:

Proof. By definition 4.1.1 the process Y_t is solvable if for all λ we can reduce equation

$$\mathcal{L}_Y f(y) = \frac{1}{2}\sigma_Y^2(y)\frac{\partial^2 f(y)}{\partial y^2} = \lambda f(y)$$

to the (scaled confluent) hypergeometric equation by a Liouville transformation, with a change of variables $y = y(x)$ independent of λ. The Liouville transformation consists of three parts: a change of variables, a multiplication by function and a gauge transformation and all these transformations commute. For example we could first apply a multiplication by function transformation, and then use only change of variables and gauge transformations. Thus we can reformulate the problem as follows: find all functions $\sigma_Y(y)$, such that there exist a function $\gamma(y)$, such that the equation

$$2\gamma^2(y)\mathcal{L}_Y f(y) = \sigma_Y^2(y)\gamma^2(y)\frac{\partial^2 f(y)}{\partial y^2} = 2\lambda\gamma^2(y)f(y) \tag{4.3.11}$$

can be reduced to the (confluent) hypergeometric equation by a change of variables (independent of λ) and a gauge transformation. The Bose invariant of equation (4.3.11) is given by

$$I(x) = \frac{1}{2}\{y, x\} - 2\lambda\gamma^2(y(x)), \tag{4.3.12}$$

with the change of variables given by

$$\frac{dy}{dx} = \sigma_Y(y)\gamma(y). \tag{4.3.13}$$

Note that $\gamma(y)$ must be independent of λ since $\sigma_Y(y)$ and $y'(x)$ are independent of λ.

By lemma 4.2.2 we know that equation (4.3.11) can be reduced to the (confluent) hypergeometric equation by a change of variables and a gauge transformation if and only if the corresponding Bose invariants are equal, that is

$$I(x) = I_{hyp}(x), \quad \text{or } I(x) = I_{confl}(x),$$

thus, combining equations (4.3.12) with (4.3.3) and (4.3.5), we have

$$\frac{1}{2}\{y, x\} - 2\lambda\gamma^2(y(x)) = \frac{Q(x)}{A^2(x)},$$

where $A(x) \in \{x, x(1-x)\}$ and $Q(x) = Q(x, \lambda)$ is some second order polynomial in x with parameters depending on λ. Note that $\{y, x\}$ and $\gamma^2(y(x))$ are independent of λ, thus there exist two polynomials $T(x), R(x) \in P_2$, independent of λ, such that

$$\begin{cases} Q(x) = T(x) - 2\lambda R(x), \\ \frac{1}{2}\{y, x\} = \frac{T(x)}{A^2(x)}, \\ \gamma^2(y(x)) = \frac{R(x)}{A^2(x)}, \quad R(x) \geq 0. \end{cases} \tag{4.3.14}$$

The last equation in the system (4.3.14) gives us the function $\gamma(y(x))$:

$$\gamma(y(x)) = \frac{\sqrt{R(x)}}{A(x)}. \tag{4.3.15}$$

By theorem 4.3.4, the solutions to equation $\frac{1}{2}\{y,x\} = \frac{T(x)}{A^2(x)}$ are given by

$$y(x) = \frac{c_3 f_1(x) + c_4 f_2(x)}{c_1 f_1(x) + c_2 f_2(x)}, \tag{4.3.16}$$

where f_1 and f_2 are linearly independent solutions to equation

$$f''(x) + \frac{T(x)}{A^2(x)} f(x) = 0. \tag{4.3.17}$$

Now we can use corollary (4.3.3), which tells us that all the solutions to equation (4.3.17) can be found in the form $F(x)/\sqrt{W(x)}$, thus

$$y(x) = \frac{c_3 F_1(x) + c_4 F_2(x)}{c_1 F_1(x) + c_2 F_2(x)} \tag{4.3.18}$$

where F_1 and F_2 are confluent hypergeometric ($A(x) = x$) or hypergeometric ($A(x) = x(1-x)$) functions.

Now we are ready to find volatility function $\sigma_Y(y)$. From the equation (4.3.13) we find that

$$\sigma_Y(Y(x)) = y'(x)\frac{1}{\gamma(y(x))}. \tag{4.3.19}$$

The derivative $y'(x)$ can be computed as

$$y'(x) = \frac{CW(x)}{(c_1 F_1(x) + c_2 F_2(x))^2},$$

thus

$$\sigma_Y(Y(x)) = y'(x)\frac{1}{\gamma(y(x))} = C\frac{CW(x)}{(c_1 F_1(x) + c_2 F_2(x))^2}\frac{A(x)}{\sqrt{R(x)}},$$

which completes the proof. □

Definition 4.3.6. We call the family of driftless processes with volatility function given by equation (4.3.1) a *hypergeometric R-family* in the case $A(x) = x(1-x)$ and a *confluent hypergeometric R-family* in the case $A(x) = x$.

We see that in the case $R(x) = A(x)$ we recover the hypergeometric and confluent hypergeometric families, which correspond to $\mathfrak{M}$(Jacobi) and $\mathfrak{M}$(CIR). In the case $R(x) \neq A(x)$ we obtain new families of processes. The next theorem shows that the (confluent) hypergeometric R-family can be obtained by stochastic transformations described from some diffusion process (in the same way as Jacobi and CIR families are generated by a single diffusion process).

Theorem 4.3.7. Second classification theorem. *Let $R(x) \in P_2$ be a second degree polynomial in x.*

(i) The confluent hypergeometric case: Assume that $R(x)$ has no zeros in $(0, \infty)$. Let $X_t = X_t^R$ be the diffusion process with dynamics

$$dX_t = (a + bX_t)\frac{X_t}{R(X_t)}dt + \frac{X_t}{\sqrt{R(X_t)}}dW_t. \tag{4.3.20}$$

Then the confluent hypergeometric R-family coincides with $\mathfrak{M}(X_t^R)$ and thus can be obtained from X_t^R by stochastic transformations. In the particular case $R(x) = A(x) = x$ we obtain the CIR family defined in (3.4.4).

(ii) The hypergeometric case: Assume that $R(x)$ has no zeros in $(0, 1)$. Let $X_t = X_t^R$ be the diffusion process with dynamics:

$$dX_t = (a + bX_t)\frac{X_t(1 - X_t)}{R(X_t)}dt + \frac{X_t(1 - X_t)}{\sqrt{R(X_t)}}dW_t. \tag{4.3.21}$$

Then the hypergeometric R-family coincides with $\mathfrak{M}(X_t^R)$ and thus can be obtained from X_t^R by stochastic transformations. In the particular case $R(x) = A(x) = x(1 - x)$ we obtain the Jacobi family defined in (3.4.5).

Proof. The generator of X_t is given by

$$\mathcal{L}_X = (a + bx)\frac{A(x)}{R(x)}\frac{\partial}{\partial x} + \frac{1}{2}\frac{A^2(x)}{R(x)}\frac{\partial^2}{\partial x^2}. \tag{4.3.22}$$

One way to prove this theorem is to check that the corresponding Bose invariants coincide. However we prove this theorem by applying stochastic transformations to the process X_t and showing that we can cover all of the processes with volatility functions given by equation (4.3.1).

First we need to find two linearly independent solutions to the "eigenfunction" equation

$$\mathcal{L}_X \varphi = \rho\varphi.$$

This equation is equivalent to

$$2(a + bx)\frac{\partial\varphi(x)}{\partial x} + A(x)\frac{\partial^2\varphi(x)}{\partial x^2} = 2\rho\frac{R(x)}{A(x)}\varphi(x).$$

By dividing both sides by $A(x)$ and making gauge transformation with gauge factor $h(x) = \sqrt{W(x)}$, $F = \varphi/h$, we arrive at the equation in canonical form:

$$\frac{\partial^2 F(x)}{\partial x^2} + \frac{Q(x) - 2\alpha R(x)}{A^2(x)}F(x) = 0.$$

By corollary 4.3.3 this equation is solved in terms of hypergeometric functions, thus $\varphi_i(x) = g(x)F_i(x)$, where F_i are (confluent) hypergeometric functions. Thus

$$Y(x) = \frac{c_1\varphi_1(x) + c_2\varphi_2(x)}{c_3\varphi_1(x) + c_4\varphi_2(x)} = \frac{c_1F_1(x) + c_2F_2(x)}{c_3F_1(x) + c_4F_2(x)},$$

and $\sigma_Y(y)$ is computed as

$$\sigma_Y(y) = Y'(x)\frac{A(x)}{\sqrt{R(x)}} = C\sqrt{A(x)}\frac{W(x)}{(c_1F_1(x) + c_2F_2(x))^2}\sqrt{\frac{A(x)}{R(x)}},$$

which ends the proof. □

Chapter 5

Applications

In this chapter we apply the theory developed in the previous chapters to several problems in Mathematical Finance. In the first section we prove that Meixner and Charlier processes (which are lattice approximations to CIR and OU) are also affine. We compute explicitly the generating function of each process and we give a method of computing the (discounted) transitional probabilities as a Fourier transform of the generating function.

Then we discuss the concept of time change, and we show how one can introduce jumps and stochastic volatility in the models constructed in chapter 3.

In the next section we use the solvable driftless process with stochastic volatility and jumps to model the stock price under the risk neutral measure and we find an explicit expression for the price of the European call option in this model (also in the case of a process on the lattice). After that we present an algorithm to price American style options with stochastic volatility and jumps.

5.1 Affine Lattice models

Let r_t be a stationary Markov process.

Definition 5.1.1. Let $\alpha, \lambda \in \mathbb{C}$. The *generating function* of the process r_t is defined as

$$G[t, r, \alpha, \lambda] = E\left(e^{\alpha r_t - \lambda \int_0^t r_s ds} | r_0 = r\right). \tag{5.1.1}$$

The *characteristic function* $F[t, r, \alpha]$ of the process r_t is instead defined

$$F[t, r, \alpha] = E\left(e^{i\alpha r_t} | r_0 = r\right). \tag{5.1.2}$$

The function $G[t, r, \alpha, \lambda]$ can be evaluated by means of the following standard result:

Lemma 5.1.2. *Let $\mathcal{L}_r$ be the Markov generator of the process r_t. The generating function $G[t, r, \alpha, \lambda]$ solves the following evolution equation*

$$-\frac{\partial G}{\partial t} + \mathcal{L}_r G = \lambda r G \tag{5.1.3}$$

with initial condition $G[0, r, \alpha, \lambda] = e^{\alpha r}$.

Definition 5.1.3. A process r_t is called *affine* if the generating function has the form

$$G[t, r, \alpha, \lambda] = e^{m(t,\alpha,\lambda)r+n(t,\alpha,\lambda)} \tag{5.1.4}$$

for some deterministic functions $m(t, \alpha, \lambda)$ and $n(t, \alpha, \lambda)$ (thus $\log(G)$ is an affine function of r).

It is a well known fact that both Ornstein-Uhlenbeck and CIR processes are affine. The following is the main theorem in this section and shows that both the Meixner and Charlier processes (see (2.2.2) and (2.2.4)) are affine and analytically solvable, by explicitly evaluating the functions $m(t, \alpha, \lambda)$ and $n(t, \alpha, \lambda)$:

Theorem 5.1.4. *Meixner and Charlier processes are affine for all values of the parameters.*

(i) In the case r_t is a Charlier process the functions m and n are given by

$$\begin{cases} m(t,\alpha,\lambda) = \alpha - \frac{1}{\lambda\delta}\log\left[1 + (1-e^{-(b-\lambda\delta)t})(\frac{b}{b-\lambda\delta}e^{\delta\alpha} - 1)\right] \\ n(t,\alpha,\lambda) = \frac{\sigma^2+2a\delta}{2\delta^2 b}\left(\log\left[1 + (1-e^{-(b-\lambda\delta)t})(\frac{b}{b-\lambda\delta}e^{\delta\alpha} - 1)\right] - \lambda\delta t\right) + \\ \qquad + \frac{\sigma^2}{2\delta^2(b-\lambda\delta)}\left((1-e^{-(b-\lambda\delta)t})(e^{-\delta\alpha} - \frac{b}{b-\lambda\delta}) + \lambda\delta t\right) \end{cases} \tag{5.1.5}$$

(ii) In the case r_t is a Meixner process the functions m and n are given by

$$\begin{cases} m(t,\alpha,\lambda) = \frac{1}{\delta}\log\left[1 + M_1 - \frac{\gamma}{\eta} + \frac{e^{\delta\alpha}-1-M_1+\frac{\gamma}{\eta}}{1+\frac{\eta}{\gamma}(1-e^{\gamma t})(e^{\delta\alpha}-1-M_1)}\right] \\ n(t,\alpha,\lambda) = \frac{a}{\delta}\left(M_1 t - \frac{1}{\eta}\log\left[1 + \frac{\eta}{\gamma}(1-e^{\gamma t})(e^{\delta\alpha} - 1 - M_1)\right]\right) \end{cases} \tag{5.1.6}$$

where $\eta = \frac{\sigma^2}{2\delta} - b$, $\gamma = \sqrt{(b+\lambda\delta)^2 + 4\lambda\delta\eta}$ and $M_1 = \frac{b+\lambda\delta+\gamma}{2\eta}$.

Proof. The generators of Charlier ($\beta = 0$)and Meixner ($\beta = 1/2$) process are given by

$$\mathcal{L} = (a - br)\nabla_\delta + \frac{1}{2}\sigma^2 r^{2\beta}\Delta_\delta. \tag{5.1.7}$$

By using (5.1.4) as an ansatz to solve the equation for the discount factor (5.1.3), we find

$$-\dot{m}r - \dot{n} + (a - br)\frac{1}{\delta}(e^{\delta m} - 1) + \frac{1}{2}\sigma^2 r^{2\beta}\frac{1}{\delta^2}(e^{\delta m} + e^{-\delta m} - 2) = \lambda r.$$

In the case $\beta = 0$ this equation splits into the two differential equations

$$\begin{cases} \dot{m} = -\frac{b}{\delta}(e^{\delta m} - 1) - \lambda \\ \dot{n} = \frac{a}{\delta}(e^{\delta m} - 1) + \frac{1}{2}\frac{\sigma^2}{\delta^2}(e^{\delta m} + e^{-\delta m} - 2) \end{cases} \tag{5.1.8}$$

while if $\beta = \frac{1}{2}$ we have the following system:

$$\begin{cases} \dot{m} = -\frac{b}{\delta}(e^{\delta m} - 1) - \lambda + \frac{1}{2}\frac{\sigma^2}{\delta^2}(e^{\delta m} + e^{-\delta m} - 2) \\ \dot{n} = \frac{a}{\delta}(e^{\delta m} - 1) \end{cases} \tag{5.1.9}$$

In both cases the initial conditions are $m(0, \alpha, \lambda) = \alpha$, $n(0, \alpha, \lambda) = 0$. We will illustrate how to solve these equation on the example of Meixner process (the case of Charlier process can be treated similarly).

Let's introduce the new function $M(t, \alpha, \lambda) = e^{\delta m(t,\alpha,\lambda)} - 1$. Multiplying the first equation in the system (5.1.9) by $\delta e^{\delta m}$ we arrive at

$$\begin{aligned} \dot{M} = \delta e^{\delta m}\dot{m} &= -b(e^{2\delta m} - e^{\delta m}) - \lambda\delta e^{\delta m} + \frac{1}{2}\frac{\sigma^2}{\delta}(e^{2\delta m} - 2e^{\delta m} + 1) \\ &= \eta M^2 - (b + \lambda\delta)M - \lambda\delta, \end{aligned} \tag{5.1.10}$$

where we denote $\eta = \frac{\sigma^2}{2\delta} - b$. The initial condition is $M(0, \alpha, \lambda) = e^{\delta\alpha} - 1$.

Equation (5.1.10) is a Ricatti equation and can be solved explicitly:

$$M(t, \alpha, \lambda) = \frac{M_1 - M_2 e^{\gamma t}\frac{M_1 - e^{\delta\alpha} - 1}{M_2 - e^{\delta\alpha} - 1}}{1 - e^{\gamma t}\frac{M_1 - e^{\delta\alpha} - 1}{M_2 - e^{\delta\alpha} - 1}}, \tag{5.1.11}$$

where $\gamma = \sqrt{(b + \lambda\delta)^2 + 4\lambda\delta\eta}$ and $M_{1,2} = \frac{b + \lambda\delta \pm \gamma}{2\eta}$ are the roots of the second order polynomial $\eta M^2 - (b + \lambda\delta)M - \lambda\delta$. One can "simplify" (5.1.11) and arrive at the formula for function $m(t, \alpha, \lambda)$ (see eq. (5.1.6)).

After the function $M(t, \alpha, \lambda)$ is found, function n can be computed using the second equation in the system (5.1.9) as the following integral

$$n(t, \alpha, \lambda) = \frac{a}{\delta}\int_0^t M(s, \alpha, \lambda)ds.$$

□

Corollary 5.1.5. *In the case of Charlier and Meixner processes the characteristic function $F[t,r,\alpha]$ can be computed explicitly and is given by*

$$F[t,r,\alpha] = e^{\bar{m}(t,\alpha)r+\bar{n}(t,\alpha)} \tag{5.1.12}$$

where functions $\bar{m}(t,\alpha)$ and $\bar{n}(t,\alpha)$ are given below.

(i) If r_t is a Charlier process, then

$$\begin{cases} \bar{m}(t,\alpha) = i\alpha - \frac{1}{\delta}\log[1+(1-e^{-bt})(e^{i\delta\alpha}-1)] \\ \bar{n}(t,\alpha) = \frac{\sigma^2}{2b\delta^2}((1+\frac{2a\delta}{\sigma^2})\log[1+(1-e^{-bt})(e^{i\delta\alpha}-1)] + (1-e^{-bt})(e^{-i\delta\alpha}-1)). \end{cases} \tag{5.1.13}$$

(ii) If r_t is a Meixner process, then

$$\begin{cases} \bar{m}(t,\alpha) = -\frac{1}{\delta}\log\left[1+\frac{\frac{2b\delta}{\sigma^2}(1-e^{-i\delta\alpha})e^{-bt}}{(1-e^{-i\delta\alpha})(1-e^{-bt})-\frac{2b\delta}{\sigma^2}}\right] \\ \bar{n}(t,\alpha) = \frac{\frac{2a}{\sigma^2}}{\frac{2b\delta}{\sigma^2}-1}\log\left[1+(e^{i\delta\alpha}-1)(1-e^{-bt})(1-\frac{\sigma^2}{2b\delta}).\right] \end{cases} \tag{5.1.14}$$

5.1.1 Computing discounted transition probabilities by Fourier transform

Let r_t be a stationary Markov process on the lattice Λ.

Definition 5.1.6. The *transition probabilities* for r_t are defined by

$$p_r(t,j,k) = \mathrm{P}\left(r_t = k\delta | r_0 = j\delta\right) = E\left(\mathbb{I}(r_t = k\delta)|r_0 = j\delta\right). \tag{5.1.15}$$

Here $\mathbb{I}(A)$ is the indicator of the set A. The *discounted transition probabilities* $q_r(t,j,k)$ are defined by

$$q_r(t,j,k) = E\left(\mathbb{I}(r_t = k\delta)e^{-\int_0^t r_s ds}|r_0 = j\delta\right) \tag{5.1.16}$$

One can easily check that the discounted transition probabilities satisfy the Chapman-Kolmogorov equation:

$$q_r(t+s,j,k) = \sum_l q_r(s,j,l)q_r(t,l,k).$$

Thus these probabilities correspond to the process r_t with killing at the rate proportional r_t.

Transitional probabilities are useful when computing discounted expectations of the form

$$E\left(f(r_T)e^{-\int_t^T r_s ds}|r_t = j\delta\right) = \sum_k f(k\delta)q_r(T-t,j,k).$$

Lemma 5.1.7. *Discounted transition probabilities* $q_r(t,j,k)$ *can be computed as the inverse Fourier transform*

$$q_r(t,j,k) = \int_0^1 e^{-2\pi i\omega k} G[t, j\delta, 2\pi i\frac{\omega}{\delta}]d\omega. \tag{5.1.17}$$

Proof. The definition of the generating function can be rewritten in the form

$$G[t, j\delta, \alpha] = \sum_{l=0}^{\infty} q_r(t,j,l)e^{\alpha l\delta}. \tag{5.1.18}$$

Thus the integral in the right side of the equation (5.1.17) is equal to

$$\int_0^1 e^{-2\pi i\omega k} \sum_{l=0}^{\infty} q_r(t,j,l)e^{2\pi i\omega l}d\omega = \sum_{l=0}^{\infty} q_r(t,j,l)\int_0^1 e^{-2\pi i\omega k}e^{2\pi i\omega l}d\omega = q_r(t,j,k)$$

since

$$\int_0^1 e^{-2\pi i\omega k}e^{2\pi i\omega l}d\omega = \delta(l-k.)$$

□

Corollary 5.1.8. *The transition probabilities* $p_r(t,j,k)$ *can be computed as the inverse Fourier transform*

$$p_r(t,j,k) = \int_0^1 e^{-2\pi i\omega k} F\left(t, j\delta, 2\pi\frac{\omega}{\delta}\right)d\omega. \tag{5.1.19}$$

Remark 5.1.9. Lemma 5.1.7 tells us that we can find the transitional probabilities as an inverse Fourier transform of a generating function. However in practice one could use slightly different argument to compute $q_r(t,j,k)$ more efficiently.

The idea is the following: the generating function $G[t, j\delta, 2\pi i\frac{\omega}{\delta}]$ is a Fourier transform of $q_r(t,j,k)$, thus we could use the information that $q_r(t,j,k)$ is a discounted transition probability to simplify the computations.

If r_s is positive (as in a Meixner process) or positive with probability close to one, then discounted transition probabilities are bounded from above by transition probabilities:

$$q_r(t,j,k) \le p_r(t,j,k).$$

Now we proceed as following: if t is not too large, then for every j there is an interval $[0, k_2]$, such that $p_r(t,j,k)$ is negligibly small (say less then some ϵ) if k is outside this interval, thus (neglecting terms of order ϵ) we can assume that $p_r(t,j,k)$ (and thus $q_r(t,j,k)$) have finite support. Then $G[t,j\delta,\alpha_n]$ is just a finite discrete Fourier transform, and thus $q_r(t,j,k)$ can be computed as an inverse finite discrete Fourier transform of the vector $G[t,j\delta,\alpha_n]$:

$$q_r(t,j,k) = \frac{1}{k_2+1}\sum_{n=0}^{k_2} G[t,j\delta,\alpha_n]e^{-2\pi i\frac{nk}{k_2+1}} \tag{5.1.20}$$

where $\alpha_n = 2\pi i\frac{n}{(k_2+1)\delta}$, $n = \{0,1,..,k_2\}$.

5.2 Stochastic time change

In this section we discuss the concept of time change, and we show how one can introduce jumps and stochastic volatility in the models constructed in chapter 3.

Definition 5.2.1. We call T_t a *time change process* if T_t is an increasing, right continuous process with $T_0 = 0$. Given a process X_t we define the *time-changed process* $\tilde{X}_t$ to be:

$$\tilde{X}_t = X_{T_t}. \tag{5.2.1}$$

The following well-known theorem (see [24], vol. II , p.277) shows that stochastic time change is an extremely useful tool in stochastic analysis:

Theorem 5.2.2. *Let X_t be a local martingale diffusion process with generator*

$$\mathcal{L}_X = \frac{1}{2}\sigma_X^2(x)\frac{\partial^2}{\partial x^2}.$$

Then for each $y \in D$ there exists on some enrichment of probability space $(\Omega,\mathcal{F},\mathrm{P}_y)$ a Brownian motion W, started at y such that the canonical process of X (a diffusion with law P_y on canonical space) can be expressed as a time-changed Brownian motion:

$$X_t = W_{\gamma_t},\quad \gamma_t = \int_0^t \sigma_X^2(X_s)ds. \tag{5.2.2}$$

From now on we adopt the following

Main assumption: the time change process T_t is independent of the process X_t.

In this case, if $\rho_t(ds)$ is the distribution of T_t, we have that

$$E_{y_0}f(\tilde{Y}_t) = \int_0^\infty [E_{y_0}f(Y_s)]\rho_t(ds). \tag{5.2.3}$$

This fact though very simple is actually the main result in this section. As the first application of formula (5.2.3) lets compute the transition density for the time-changed process $\tilde{Y}_t$.

Let X_t be one of the processes described in chapter 2, thus its probability density can be expressed as

$$p_X(t, x_0, x_1) = \sum_{n=0}^{\infty} e^{\lambda_n t} \psi_n(x_0)\psi_n(x_1). \tag{5.2.4}$$

Let $Y_t = (Y(X_t), \mathrm{P}^h)$ be related to X_t by a stochastic transformation $\{\rho, h, Y\}$. Then as we know from lemma 3.3.1 the probability density for Y is computed as

$$p_Y(t, y_0, y_1) = \frac{e^{-\rho t}}{h(x_0)h(x_1)} p_X(t, x_0, x_1) = \frac{1}{h(x_0)h(x_1)} \sum_{n=0}^{\infty} e^{-(\rho-\lambda_n)t} \psi_n(x_0)\psi_n(x_1), \tag{5.2.5}$$

where as usual $y_i = Y(x_i)$. Note that the time and space variables are conveniently separated in equation (5.2.5). Now using formula (5.2.3) we obtain the expression for $p_{\tilde{Y}}$:

$$\begin{aligned} p_{\tilde{Y}}(t, y_0, y_1) &= \int_0^{\infty} p_Y(s, y_0, y_1)\rho_t(ds) \\ &= \frac{1}{h(x_0)h(x_1)} \sum_{n=0}^{\infty} \left[\int_0^{\infty} e^{-(\rho-\lambda_n)t} \rho_t(ds) \right] \psi_n(x_0)\psi_n(x_1) \\ &= \frac{1}{h(x_0)h(x_1)} \sum_{n=0}^{\infty} L(\rho - \lambda_n, t)\psi_n(x_0)\psi_n(x_1), \end{aligned}$$

where

$$L(\lambda, t) = E e^{-\lambda T_t} \tag{5.2.6}$$

is the *Laplace transform* of random variable T_t. Thus we have proved the following

Lemma 5.2.3. *Let the process X_t be associated with a family of orthogonal polynomials $\psi_n(x)$ and assume that Y_t is obtained by a stochastic transformation $\{\rho, h, Y\}$ from the process X_t. The probability density for the process $\tilde{Y}_t = Y_{T_t}$ is given by*

$$p_{\tilde{Y}}(t, y_0, y_1) = \frac{1}{h(x_0)h(x_1)} \sum_{n=0}^{\infty} L(\rho - \lambda_n, t)\psi_n(x_0)\psi_n(x_1). \tag{5.2.7}$$

Note that the process $\tilde{Y}$ is solvable if we can compute the Laplace transform of T_t. Thus we restrict ourselves only to time change processes for which the Laplace transform can be computed explicitly, the two main examples (increasing Levy process and integral of an affine process) will be discussed in the next two sections.

5.2.1 Increasing jump processes: subordinators

Let $(\Omega, \mathcal{F}_t, \mathrm{P})$ be a probability space and let γ_t be the right-continuous increasing process, started from zero $\gamma_0 = 0$ and with values in $(0, \infty)$.

Definition 5.2.4. γ_t is called a subordinator if it has independent and time-homogeneous increments: for any $0 < s < t$ we have that $\gamma_t - \gamma_s$ is independent of $\mathcal{F}_s$ and has the same distribution as γ_{t-s}.

Remark 5.2.5. Subordinators can be equivalently defined as increasing Levy processes.

Let $\rho_t(dy)$ be the one-dimensional distribution of γ_t: $\rho_t(dy) = \mathrm{P}(\gamma_t \in dy)$. Then ρ_t generates a *convolution semigroup*:

$$P_t f(x) = \int\limits_{[0,\infty)} f(x+y)\rho_t(dy),$$

for all bounded Borel measurable $f(x)$.

Note that since γ_t has independent increments, the Laplace transform of γ_t has the following multiplicative property:

$$E\left(e^{-\lambda\gamma_{t+s}}\right) = E\left(e^{-\lambda(\gamma_{t+s}-\gamma_t)}e^{-\lambda\gamma_t}\right) = E\left(e^{-\lambda\gamma_t}\right) E\left(e^{-\lambda\gamma_s}\right)$$

Thus we can express the Laplace transform of γ_t as $\exp(-t\Phi(\lambda))$, where the function $\Phi : [0,\infty) \to [0,\infty)$ is called the *Laplace exponent* of γ_t. The following well known result gives a way to compute $\Phi(\lambda)$ and characterizes subordinators:

Theorem 5.2.6. Levy, Khintchine

(i) There exist a unique number $d \geq 0$ and a unique measure $\Pi(dx)$ on $(0,\infty)$, satisfying $\int(1 \wedge x)\Pi(dx) < \infty$, such that for every $\lambda \geq 0$

$$\Phi(\lambda) = d\lambda + \int_{(0,\infty)} \left(1 - e^{-\lambda x}\right)\Pi(dx). \tag{5.2.8}$$

(ii) The converse is true: any function $\Phi(\lambda)$ which can be represented in the form (5.2.8), is a Laplace exponent of some subordinator.

The measure $\Pi(dx)$ is called the *Levy measure* for the process γ_t. It's probabilistic meaning is the following: $\Pi(A)dt$ is the expected number of jumps $\Delta\gamma_t$ of the size $\Delta\gamma_t \in A$ in the time interval $(t, t+dt)$.

When $\Pi((0,\infty)) < \infty$, γ_t can be modelled as *the compound Poisson process*:

$$\gamma_t = \sum_{n=0}^{N_t} \xi_n$$

where ξ_n are independent random variables with distribution given by $\mathrm{P}(\xi < x) = \Pi((0,x))/\Pi((0,\infty))$ and N_t is a Poisson process with parameter $\Lambda = \Pi((0,\infty))$.

When $\Pi((0,\infty)) = \infty$ we have an infinite number of jumps in any interval $(t, t+dt)$.

Next we will present some examples of subordinators, which we will use in later sections to model the jump component of a stochastic time change process:

- The poisson process with intensity Λ is the simplest example of the subordinator: it corresponds to Laplace exponent

$$\Phi(\lambda) = \Lambda(1 - e^{-\lambda}),$$

  with the Levy measure $\Pi(dx) = \Lambda\delta_1$. The one-dimensional distribution is just the Poisson distribution.

- The stable processes:

$$\Phi(\lambda) = \lambda^\alpha = \frac{\alpha}{\Gamma(\alpha)} \int_0^\infty \left(1 - e^{-\lambda x}\right) x^{-1-\alpha} dx, \tag{5.2.9}$$

  where the requirement $\int (1 \wedge x)\Pi(dx) < \infty$ gives $\alpha \in (0,1)$. The boundary case $\alpha = 1$ is degenerate since it corresponds to $\gamma_t = t$. The explicit formula for the one dimensional distribution is known only in the case $\alpha = 1/2$, in which case $\gamma_t = \tau_t$ is the first passage process of a linear Brownian motion:

$$\tau_t = \int \{s \geq 0 : \quad B_s > t\}.$$

  Due to the scaling property of Brownian motion and the reflection principle one can compute the distribution of τ_1:

$$\mathrm{P}(\tau_1 < t) = \frac{1}{\sqrt{2\pi}} \int_0^t s^{-3/2} e^{-1/2s} ds.$$

- The Gamma process:

$$\rho_t^{\Gamma}(ds) = \frac{\left(\frac{\mu}{\nu}\right)^{\frac{\mu^2 t}{\nu}}}{\Gamma(\frac{\mu^2 t}{\nu})} s^{\frac{\mu^2 t}{\nu}-1} e^{-\frac{\mu}{\nu}s} ds. \tag{5.2.10}$$

where μ is the mean rate and ν is variance rate. In this case the Laplace exponent is

$$\Phi(\lambda) = \frac{\mu^2}{\nu}\ln(1+\frac{\lambda\mu}{\nu}), \tag{5.2.11}$$

and the Levy measure is given by

$$\Pi(dx) = \frac{\mu^2}{\nu}\frac{e^{-\frac{\mu}{\nu}s}}{s}ds$$

5.2.2 Absolutely continuous time change processes: affine models

In this section we consider an important example of time change process T_t^c of the form

$$T_t^c = \int_0^t r_s ds \tag{5.2.12}$$

where r_s is a positive stochastic process. Clearly T_t^c is an increasing process, absolutely continuous and it does not explode if r_s does not explode. It follows from the theorem 5.2.2 that using T_t^c as a time change process is equivalent to introducing the stochastic volatility: if X_t is a brownian motion W_t and $\tilde{X}_t = W_{T_t^c}$ is a time change of X_t, then

$$d\tilde{X}_t = dW_{T_t^c} = \sqrt{r_t}dW_t$$

thus the process r_t can be considered as a square of the volatility of the process $\tilde{X}_t$.

The Laplace transform of the process T_t^c in this case is the following expectation:

$$L_{T^c}(\lambda, t) = E\left[e^{-\lambda T_t^c}\right] = Ee^{-\lambda\int_0^t r_s ds}. \tag{5.2.13}$$

There is a vast literature on the processes r_s for which the expectation in the equation (5.2.13) can be computed explicitly (one of the reasons for that is that this expectation can be considered as a price of the zero coupon bond with the short rate process given by λr_s). The well known examples of these processes are the so called *affine processes* (see section 5.1).

The well known examples of the affine processes are the Ornstein-Uhlenbeck and the CIR process (as well as their lattice approximations given by Charlier and Meixner processes described above in section

5.1). But since the Ornstein-Uhlenbeck (and Charlier) process can attain negative values with non-zero probability, we will use only CIR (or Meixner) process for the time change modelling.

In section 5.1 we already computed the generating function for the Meixner process. Now let r_s be a CIR process with generator

$$\mathcal{L} = (c - dr)\nabla + \frac{1}{2}\zeta^2 r\Delta.$$

It is a classical result that in this case the functions $m(t) = (t, c, d, \zeta, \lambda)$ and $n(t) = n(t, c, d, \zeta, \lambda)$ can be computed as follows:

$$n(t) = \frac{2c}{\zeta_0^2}\log\left(\frac{\gamma\exp(\frac{1}{2}dt)}{\gamma\cosh(\gamma t) + \frac{1}{2}d\sinh(\gamma t)}\right), \qquad m(t) = \lambda\frac{\sinh(\gamma t)}{\gamma\cosh(\gamma t) + \frac{1}{2}d\sinh(\gamma t)} \tag{5.2.14}$$

where $\gamma = \frac{1}{2}\sqrt{d^2 + 2\lambda\zeta^2}$.

5.2.3 Unifying the three models

So far we have described the two possible ways to define a time change process in such a way to guarantee the analytical tractability of the time changed process $\tilde{Y}_t$: one of them is to use Levy subordinators, the other is to use affine processes (both continuous and processes on the lattice). The third obvious choice is to use a deterministic time change:

$$T_t^d = f(t) \tag{5.2.15}$$

for some increasing function f.

The following simple observation lets us to combine all three time change processes together: If we have two independent random variables ξ and η, then

$$Ee^{-\lambda(\xi+\eta)} = Ee^{-\lambda\xi}Ee^{-\lambda\eta}. \tag{5.2.16}$$

Thus we can take all three versions of the time change processes: deterministic time change $T_t^d = f(t)$, Levy subordinator $T_t^j = \gamma_t$ and the absolutely continuous process $T_t^c = \int_0^t r_s ds$ with r_t independent of γ_t and form a time change process:

$$T_t = T_t^d + T_t^j + T_t^c. \tag{5.2.17}$$

The Laplace transform for this process is computed as the product of Laplace transforms of each component:

$$L_T(\lambda, t) = e^{-\lambda f(t)} L_{T^j}(\lambda, t) L_{T^c}(\lambda, t). \tag{5.2.18}$$

This type of time change gives us the processes with stochastic volatility, state dependent volatility and jumps at the same time.

5.3 Pricing call options with stochastic volatility and jumps

As the first application of the above theory, in this section we use the driftless processes with stochastic volatility and jumps to model the stock price. Using the expression of transitional density in orthogonal polynomials given in lemma 5.2.3 and the Meyer-Tanaka formula we give an explicit expressions for the prices of the European call option (as expansion in orthogonal polynomials).

Definition 5.3.1. The prices of primary securities in the spot market are given by:

- savings account $B_t = 1$ (we assume for simplicity that the interest rate is 0, all the computations can be easily extended to the more general case $B_t = e^{rt}$).
- stock price S_t is given by a positive $\mathbb{Q}$-local martingale $\tilde{Y}_t$.

A call option written on a stock S is a financial security that gives it's holder the right (but not the obligation) to buy the underlying stock on a prescribed date T (maturity or an expiry date) for a prescribed price K (strike price) (see [15], p.9). In other words, a European call option is defined as financial instrument $\mathcal{X}$ with payoff function

$$\mathcal{X}_T = (S_T - K)^+. \tag{5.3.1}$$

Note that the exercise of the European call option is possible only at maturity $t = T$.

The *fundamental theorem of Finance* tells us that the spot market defined above is arbitrage free and the arbitrage price of any attainable contingent claim $\mathcal{X}$ is given by the following expectation with respect to martingale measure $\mathbb{Q}$

$$\pi_t(\mathcal{X}) = B_t E^{\mathbb{Q}}(\mathcal{X}_T B_T^{-1} | F_t) = E^{\mathbb{Q}}(\mathcal{X}_T | F_t). \tag{5.3.2}$$

In the following sections we will present an algorithm to compute efficiently the prices of a call option in the case that $\tilde{Y}_t$ is modelled by process with stochastic volatility and jumps constructed in section 5.2.3 (or its lattice approximation Y_t^δ).

5.3.1 Computation of the price of the call option for diffusion process with state-dependent volatility

Let the driftless process Y_t be the price of the stock under the risk neutral measure. Define function $C_Y(t, y_0, K)$ as the price of the call option:

$$C_Y(t, y_0, K) = E^{\mathrm{P}^h}\left((Y_t - K)^+ | Y_0 = y_0\right).$$

The following theorem gives an expression for $C_Y(t, y_0, K)$:

Theorem 5.3.2. *Let X_t be a diffusion process and $Y_t = (Y(X_t), \mathrm{P}^h)$ be a driftless process related to X_t by a stochastic transformation $\{\rho, h, Y\}$. Then*

$$\begin{aligned} C_Y(t, y_0, K) &= (y_0 - K)^+ + \\ &+ \frac{1}{2}\sigma_Y^2(K)\frac{m_Y(K)}{h(x_0)h(k)}\left[G_X(\rho, x_0, k) - \int_t^\infty e^{-\rho s} p_X(s, x_0, k)ds\right], \end{aligned} \tag{5.3.3}$$

where $x_0 = X(y_0) = Y^{-1}(y_0)$ and $k = X(K)$.

Proof. The starting point is the Meyer-Tanaka formula (analog of the Ito formula for the convex function $(y-K)^+$):

$$(Y_t - K)^+ = (Y_0 - K)^+ + \int_0^t I\{Y_s > K\}dY_s + \frac{1}{2}L_t^K(Y), \tag{5.3.4}$$

where $L_t^K(Y)$ is the local time of the process Y_t at the point K. Taking expectation on the both sides of equation (5.3.4) we find:

$$\begin{aligned} E^{\mathrm{P}^h}((Y_t - K)^+ | Y_0 = y_0) &= (y_0 - K)^+ + E^{\mathrm{P}^h}\int_0^t I\{Y_s > K\}dY_s + \frac{1}{2}E^{\mathrm{P}^h}(L_t^K(Y)|Y_0 = y_0) \\ &= (y_0 - K)^+ + \frac{1}{2}\sigma_Y^2(K)m_Y(K)\int_0^t p_Y(s, y_0, K)ds, \end{aligned}$$

where we used the property that $\int_0^t I\{Y_s > K\}dY_s$ is a P^h-martingale and thus its expectation is zero and the fact that the expectation of the local time can be computed as:

$$E^{\mathrm{P}^h}(L_t^y(Y)|Y_0 = y_0) = \sigma_Y^2(y)\int_0^t p_Y(s, y_0, y)m_Y(y)ds.$$

Next using the lemma 3.3.1 we express the probability density of Y through the probability density of X as

$$p_Y(s,y_0,K)=\frac{e^{-\rho s}}{h(x_0)h(k)}p_X(s,x_0,k),\quad x_0=X(y_0),\quad k=X(K).$$

Thus we obtain

$$\begin{aligned}E^{P^h}(L_t^K(Y)|Y_0=y_0)&=\sigma_Y^2(K)\frac{m_Y(K)}{h(x_0)h(k)}\int_0^t e^{-\rho s}p_X(s,x_0,k)ds\\&=\sigma_Y^2(K)\frac{m_Y(K)}{h(x_0)h(k)}\left[\int_0^\infty e^{-\rho s}p_X(s,x_0,k)ds-\int_t^\infty e^{-\rho s}p_X(s,x_0,k)ds\right]\\&=\sigma_Y^2(K)\frac{m_Y(K)}{h(x_0)h(k)}\left[G_X(\rho,x_0,k)-\int_t^\infty e^{-\rho s}p_X(s,x_0,k)ds\right],\end{aligned}$$

which ends the proof. □

Remark 5.3.3. The informal way of proving the theorem above is to write the Ito formula:

$$(Y_t-K)^+=(Y_0-K)^++\int_0^t I\{Y_s>K\}dY_s+\frac{1}{2}\int_0^t\delta(Y_s-K)\sigma_Y^2(Y_s)ds,$$

then again take expectations from the both sides. The integral $\int_0^t\delta(Y_s-K)\sigma_Y^2(Y_s)ds$ can be rewritten as $\int_0^t\delta(Y_s-K)\sigma_Y^2(K)ds$, and we have:

$$E^{P^h}\int_0^t\delta(Y_s-K)\sigma_Y^2(K)ds=\sigma_Y^2(K)\int_0^t E^{P^h}(\delta(Y_s-K))ds=\sigma_Y^2(K)m_Y(K)\int_0^t p_Y(s,y_0,K)ds.$$

As we see later, this idea can also be applied to the case of birth and death martingales constructed in (3.5).

5.3.2 Computation of the price of the call option for processes with state dependent volatility, stochastic volatility and jumps

Now we assume that the price of the stock is modelled by a driftless process with stochastic volatility and jumps $Y_{T_t}=\tilde{Y}_t$, and Y_t is related by a stochastic canonical transformation $\{\rho,h,Y\}$ to one of the processes described in chapter 2.

Lemma 5.3.4. *The price of the call option with an underlying $\tilde{Y}$ can be computed as:*

$$\begin{aligned}C_{\tilde{Y}}(t,y_0,K)&=(y_0-K)^++\\&+\frac{1}{2}\sigma_Y^2(K)\frac{m_Y(K)}{h(x_0)h(k)}\left[G_X(\rho,x_0,k)-\sum_{n=0}^\infty\frac{L(t,\rho-\lambda_n)}{\rho-\lambda_n}\psi_n(x_0)\psi_n(k)\right]\end{aligned}\tag{5.3.5}$$

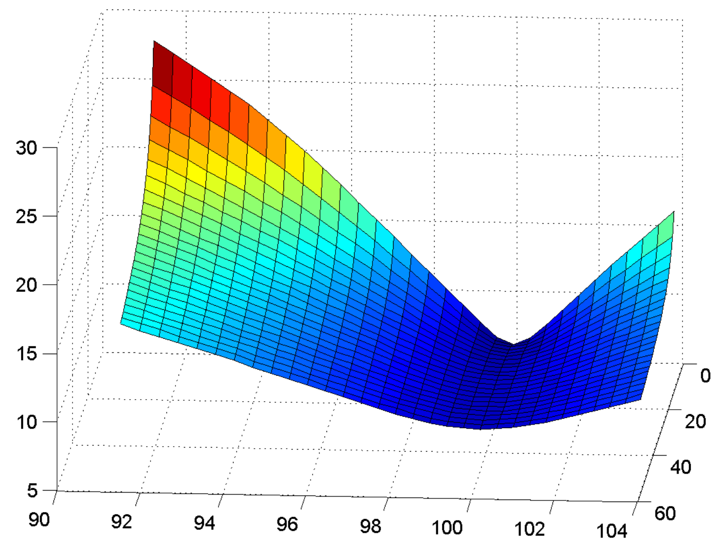

Figure 5.1: The example of an implied volatility surface in our model (note the characteristic form of ”volatility smile”, observed in market data).

Proof. At this point we will need the expression of $p_X(t, x_0, x_1)$ in terms of orthogonal basis given by the formula (5.2.4). This lets us to express the integral $\int_t^\infty e^{-\rho s} p_X(s, x_0, k)ds$ as follows:

$$\int_t^\infty e^{-\rho s} p_X(s, x_0, k)ds = \sum_{n=0}^\infty \int_t^\infty e^{-(\rho-\lambda_n)s} ds \psi_n(x_0)\psi_n(x_1) = \sum_{n=0}^\infty \frac{e^{-(\rho-\lambda_n)t}}{\rho-\lambda_n}\psi_n(x_0)\psi_n(x_1).$$

Thus using formula (5.3.3) we find that

$$C_Y(t, y_0, K) = E^{\mathrm{P}^h}((Y_t - K)^+ | Y_0 = y_0) = (y_0 - K)^+ + +\frac{1}{2}\sigma_Y^2(K)\frac{m_Y(K)}{h(x_0)h(k)}\left[G_X(\rho, x_0, k) - \sum_{n=0}^\infty \frac{e^{-t(\rho-\lambda_n)}}{\rho-\lambda_n}\psi_n(x_0)\psi_n(k)\right] \tag{5.3.6}$$

To find $C_{\tilde{Y}}(t, y_0, K) = E^{\mathrm{P}^h}((Y_{T_t} - K)^+ | Y_0 = y_0)$ we use the formula (5.2.3), thus

$$C_{\tilde{Y}}(t, y_0, K) = \int C_Y(s, y_0, K)\rho_t(ds) = (y_0 - K)^+ + +\frac{1}{2}\sigma_Y^2(K)\frac{m_Y(K)}{h(x_0)h(k)}\left[G_X(\rho, x_0, k) - \sum_{n=0}^\infty \frac{1}{\rho-\lambda_n}\left(\int e^{-s(\rho-\lambda_n)}\rho_t(ds)\right)\psi_n(x_0)\psi_n(k)\right].$$

Note that $\int e^{-s(\rho-\lambda_n)}\rho_t(ds) = L(t, \rho - \lambda_n)$, which ends the proof. □

Formula (5.3.6) can be used to efficiently compute the price of the call options. As an example we computed the implied volatility surface (see figure (5.1)).

5.3.3 Price of a call option for the process on a lattice

It is interesting to note that analog of the formula (5.3.3) can be proved in the case of birth and death martingales considered in section (3.5). As we have seen before, the birth and death processes behave a lot like diffusion – the main similarity is probably because both diffusion and birth and death processes can have no jumps, in the sense that X_t can go in the state x_1 starting at x_0 only through all the states in between x_0 and x_1. This is the main idea in the next theorem.

Let's write the generator of the birth and death process X_t^δ in the form

$$\mathcal{L}_X^\delta f(x) = a(x)f(x-\delta) - (a(x)+d(x))f(x) + d(x)f(x+\delta). \tag{5.3.7}$$

Let Y_t^δ be the process $(Y(X_t^\delta), \mathrm{P}^h)$ stopped at the boundary. Assume also that c_1 and c_2 are positive, then Y_t^δ is a bounded martingale and the following result is true:

Lemma 5.3.5. *Let* $x = X(k) = Y^{-1}(k)$ *and* $x_0 = X(y_0)$*. Then*

$$\begin{aligned} C_{Y^\delta}(t, y_0, K) &= E^{\mathrm{P}^h}\left((Y_t^\delta - K)^+ | Y_0^\delta = y_0\right) \qquad (5.3.8)\\ &= (y_0 - K)^+ + d(k)\frac{h(k+\delta)}{h(x_0)}(Y(k+\delta) - Y(k)) \int\limits_0^t e^{-\rho s} p_{X^\delta}(s, x_0, k) ds. \end{aligned}$$

Proof. We start with the *Dynkin formula* for the function $(y-K)^+$:

$$E^{\mathrm{P}^h}((Y_t^\delta - K)^+ | Y_0^\delta = y_0) = (y_0 - K)^+ + E^{\mathrm{P}^h}\left(\int\limits_0^t g(Y_s^\delta) ds | Y_0^\delta = y_0\right), \tag{5.3.9}$$

where

$$g(Y(x)) = \mathcal{L}_{X^h}^\delta (Y(x) - Y(k))^+ = \left(\frac{1}{h}\mathcal{L}_X^\delta h - \rho\right)(Y(x) - Y(k))^+.$$

Since $\mathcal{L}_{X^h}^\delta Y(x) = 0$, it is easy to check that

$$g(y) = (Y(k+\delta) - Y(k)) d(k) \frac{h(k+\delta)}{h(k)} \delta(x-k), \tag{5.3.10}$$

where $\delta(x-k)=1$ if $x=k$ and zero otherwise. Thus we have

$$
\begin{aligned}
&E^{\mathrm{P}^h}\left(\int_0^t g(Y_s^\delta)ds|Y_0^\delta=y_0\right)\\
&=(Y(k+\delta)-Y(k))d(k)\frac{h(k+\delta)}{h(k)}\int_0^t E^{\mathrm{P}}\left(e^{-\rho s}\frac{h(X_s^\delta)}{h(x_0)}\delta(X_s^\delta-k)|X_0^\delta=x_0\right)ds\\
&=(Y(k+\delta)-Y(k))d(k)\frac{h(k+\delta)}{h(k)}\frac{h(k)}{h(x_0)}\int_0^t e^{-\rho s}E^{\mathrm{P}}(\delta(X_s^\delta-k)|X_0^\delta=x_0)ds\\
&=(Y(k+\delta)-Y(k))d(k)\frac{h(k+\delta)}{h(x_0)}\int_0^t e^{-\rho s}p_{X^\delta}(s,x_0,k)ds.
\end{aligned}
$$

□

Remark 5.3.6. Notice, that for Charlier, Meixner and Hahn processes (see chapter 2), using the expansion of $p_{X^\delta}(s,x_0,k)$ in orthogonal polynomials, one could also express the integral $\int_0^t e^{-\rho s}p_{X^\delta}(s,x_0,k)ds$ in the formula (5.3.8) as an expansion in orthogonal polynomials. Then one could make a time change and arrive at the analog of lemma (5.3.4), when the process $\tilde{Y}_t$ is a process on the lattice, with stochastic volatility and jumps.

5.4 Pricing American options with stochastic volatility and jumps

Lets consider the *reward function* $g(t,y):[0,T]\times\mathbb{R}_+\to\mathbb{R}$. Define M_t^T the set of all Markov stopping times τ (with respect to initial filtration $\mathcal{F}_t$), such that $t\le\tau\le T$ a.s.

Definition 5.4.1. An *American contingent claim* $\mathcal{X}^a$ with the reward function $g(t,y)$ is a financial instrument consisting of

(i) an expiry date T

(ii) the selection of stopping times $\tau\in M_0^T$

(iii) the payoff $\mathcal{X}_\tau^a=g(\tau,S_\tau)$ on exercise.

Remark 5.4.2. Note the major difference between European and American claims is that the former can be exercised only at expiry $t=T$ while the latter can be exercised at any time $t\in[0,T]$.

The following theorem gives the expression for the price of an American claim in terms of expectation w.r.t the martingale measure (see [15], theorem 8.1.1):

Theorem 5.4.3. *The arbitrage price at time t of an American claim with reward function g equals*

$$\pi_t(\mathcal{X}^a) = \textit{ess sup}_{\tau \in M_t^T} E^{\mathbb{Q}}\left(g(\tau, S_\tau)|\mathcal{F}_t\right). \tag{5.4.1}$$

If the process S_t is Markov, there exists a value function $C^a(t,y)$, such that $\pi_t(\mathcal{X}^a) = C^a(t, S_t)$. To solve the problem of pricing an American option one needs to find the function $C^a(t,S)$ and the optimal stopping time τ^*. From now on we assume that the payoff function $g(t,y)$ is of the form $e^{-\lambda t}g(y)$.

It is known (see [23]) that the function $C^a(t,y)$ can be characterized as the least λ-excessive majorant for the function $g(y)$, that is function C^a is the least function among all functions $F(t,y)$, such that $g(y) \leq F(t,y)$ for all $t \in [0,T]$ and for all $0 \leq t \leq t + dt \leq T$

$$e^{-\lambda dt} P_S(dt) F(t,y) \leq F(t,y),$$

where P_S is the probability semigroup for the process $S = Y$. Function $C^a(t,y)$ can be found as a solution to a *free boundary problem* (also called *Stephan's problem*). The mathematical formulation is the following: let's define the two regions $C^T = \{(t,y) \in [0,T] \times D_Y : \; C^a(t,y) > g(y)\}$ and $S^T = \{(t,y) \in [0,T] \times D_Y : \; C^a(t,y) = g(y)\}$. Since the function $C^a(t,y) \geq g(y)$ we always have $S^T \bigcup C^T = [0,T] \times D_Y$.

Define the Markov stopping time $\tau_0^T = \inf\{0 \leq s \leq T : (s, S_s) \in S^T\}$. Then in "typical" situations the moment τ_0^T is the optimal stopping moment. The region S^T is called the *stopping region* and C^T *continuation region*. Function $C^a(t,S)$ satisfies the following equation

$$\begin{cases} \frac{\partial C^a(t,y)}{\partial t} - \lambda C^a(t,y) + \mathcal{L}_Y C^a(t,y) = 0, & (t,y) \in C^T \\ C^a(t,y) = g(y), & (t,y) \in S^T. \end{cases} \tag{5.4.2}$$

Function $C^a(t,S)$ can be found as a solution to the equation (5.4.2), with free boundary conditions that the values of $C^a(t,y)$ and $\frac{\partial C^a(t,y)}{\partial y}$ coincide with $g(y)$ and $g'(y)$ on the boundary of the region S^T. Note that the regions C^T and S^T are not known a priori and must be determined along with the value function $C^a(t,y)$.

In the case of American call option with payoff function given by $g(y) = (y - K)^+$ the regions C^T and S^T can be described more explicitly: there exists a nonincreasing *boundary function* $y^*(t) : \; [0,T] \in D$, such that

$$C^T = \{(t,y) \in [0,T] \times D_Y : \ y < y^*(t)\}, \tag{5.4.3}$$

$$S^T = \{(t,y) \in [0,T] \times D_Y : \ y \geq y^*(t)\}. \tag{5.4.4}$$

If $\lambda = 0$, then it can be proved that $y^*(t) = \infty$, which means that the optimal stopping time $\tau_0^T = T$ a.s. and the price of American call option coincides with the price of European call option.

We are going to use the following algorithm to approximate the price of an American option: instead of allowing exercise at any time $t \in [0,T]$, we will allow exercise only on a finite set of points $0 = t_0 < \Delta < \cdots < n\Delta = T$ (for n large enough). Thus the set of stopping times M_t^T is replaced in this case with the set $M_t^T(\Delta)$ of stopping times τ, such that $\tau \in [t,T]$ and $\tau = i\Delta$ for some i a.s. It is natural to expect that function $C^a(t,S)$ is in some sense "close" to the function

$$C^a_\Delta(t,y) = \text{ess sup}_{\tau \in M_t^T(\Delta)} E^{\mathrm{Q}}\left(e^{-\lambda(\tau - t)} g(Y_\tau) | Y_t = y\right). \tag{5.4.5}$$

We see that this problem is now the optimal stopping problem in discrete time. Function $C^a_\Delta(t,y)$ can be found recursively (see [23]):

$$C^a_\Delta(t,y) = \max\{g(y), e^{-\beta\Delta} E^{\mathrm{Q}}\left(C^a_\Delta(t+\Delta, Y_{t+\Delta}) | Y_t = y\right)\} \tag{5.4.6}$$

with final time condition $C^a_\Delta(T,y) = g(y)$.

Now we will specify our model. For the process X_t we choose the Hahn process on the finite lattice $D_X = \{0, \delta, 2\delta, \ldots, N\delta\}$ (see chapter 2), which is defined by its Markov generator

$$\mathcal{L}_X = (a - bx)\nabla_\delta + \frac{1}{2}\sigma^2 x(1-x)\Delta_\delta.$$

When δ is small this process is a lattice approximation to Jacobi process.

Next we perform the discrete version of the stochastic transformation (see section (3.5): first we need to find two linearly independent solutions φ_1, φ_2 to the finite difference equation

$$\mathcal{L}_X \varphi(x) = \rho\varphi(x), x \in \text{int}(D). \tag{5.4.7}$$

To find these solutions one would specify the values of φ_1 and φ_2 at two points in the interior of D (such that these values are linearly independent of course) and then compute all the other values of φ_i iteratively using equation (5.4.7).

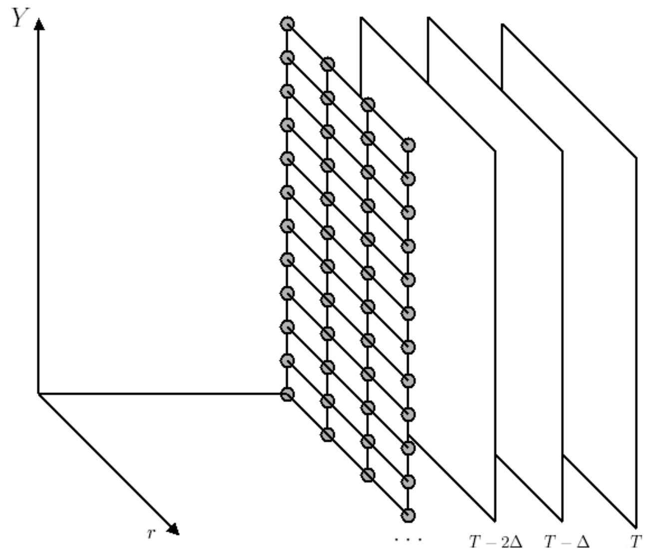

Figure 5.2: Two dimensional lattice for pricing American option with stochastic volatility and jumps.

Following theorem (3.5.1) we fix constants $c_1, \ldots, c_4$ ($c_1 > 0$ and $c_2 > 0$ and $c_1c_4 - c_2c_3 \neq 0$) and define functions

$$h(x) = c_1\varphi_1(x) + c_2\varphi_2(x), \quad Y(x) = \frac{c_3\varphi_1(x) + c_4\varphi_2(x)}{h(x)}.$$

Then the process $Y_t = (Y(X_t), \mathrm{P}^h)$, stopped at the boundary, is a bounded martingale under the measure P^h (which is the h-transform of measure P).

Remark 5.4.4. Note that the process Y_t is a martingale only before it hits the boundary, but by choosing the right parameters we can make sure that for fixed T the probability of hitting the boundary (starting at Y_0 at $t = 0$) before $t = T$ is negligibly small.

Now we use the concept of stochastic time change to obtain a martingale $\tilde{Y}_t$ with "stochastic volatility and jumps". We model the time change process as

$$T_t = f(t) + \int_0^t r_s ds + \gamma_t, \tag{5.4.8}$$

where $f(t)$ is some increasing deterministic function, r_s is a positive affine process and γ_t is a Levy subordinator. For the affine process r_s we choose a Meixner process (independent of X_s). Define the process $\tilde{Y}_t = Y_{T_t}$.

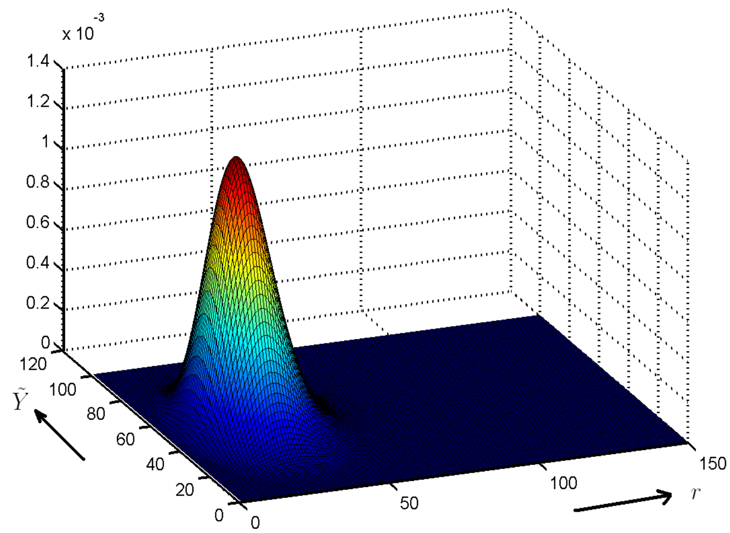

Figure 5.3: An example of the matrix of transition probabilities for the process $(\tilde{Y}, r)$.

Note that the process $\tilde{Y}_t$ is not a Markov process (its dynamics depends on a stochastic process r_s), but the two-dimensional process $(r_t, \tilde{Y}_t)$ is Markov. Thus the price of an American option also depends on variable r:

$$\pi_t(\mathcal{X}^a) = C^a_\Delta(t, y, r) = \text{ess sup}_{\tau \in M^T_t(\Delta)} E^{\mathrm{Q}} \left(e^{-\lambda(\tau - t)} g(\tilde{Y}_\tau) | \tilde{Y}_t = y, r_t = r \right).$$

To price an American style option using the recursive procedure described by equation (5.4.6) we use the transitional probabilities of the two dimensional Markov process $(r_t, \tilde{Y}_t)$:

$$p_{\tilde{Y},r}(t, y_0, y_1; r_0, r_1) = \mathrm{P}^h(\tilde{Y}_t = y_1, r_t = r_1 | \tilde{Y}_0 = y_0, r_0).$$

The following lemma gives a numerically efficient method to compute these probabilities:

Lemma 5.4.5. *The transitional probabilities of the Markov process* $(r_t, \tilde{Y}_t)$ *can be computed as*

$$p_{\tilde{Y},r}(t, y_0, y_1; r_0, r_1) = m(x_1) \frac{h(x_1)}{h(x_0)} \sum_{n=0}^{\infty} l_n(t, r_0, r_1) \psi_n(x_0) \psi_n(x_1) \tag{5.4.9}$$

where the factor $l_n(t, r_0, r_1)$ *is given by*

$$l_n(t, r_0, r_1)) = e^{-(\rho - \lambda_n) f(t)} L_\gamma(\rho - \lambda_n, t) q_r(\rho - \lambda_n, t, r_0, r_1), \tag{5.4.10}$$

and $q_r(\rho - \lambda_n, t, r_0, r_1)$ *are the discounted transition probabilities of the process* $(\rho - \lambda_n) r_s$*:*

$$q_r(\rho - \lambda_n, t, r_0, r_1) = E\left(e^{-(\rho - \lambda_n) \int_0^t r_s ds} | r_t = r_1, r_0 \right). \tag{5.4.11}$$

Proof.

$$p_{\tilde{Y},r}(t,y_0,y_1;r_0,r_1) = E^{\mathrm{P}^h}(\delta(\tilde{Y}_t - y_1)\delta(r_t - r_1)|\tilde{Y}_{t=0} = y_0, r_0)$$
$$= m(x_1)\frac{h(x_1)}{h(x_0)}\sum_{n=0}^{\infty} E^{\mathrm{P}^h}\left(e^{-(\rho-\lambda_n)T_t}|r_t = r_1, r_0\right)\psi_n(x_0)\psi_n(x_1),$$

where the expression $E^{\mathrm{P}^h}\left(\exp(-(\rho-\lambda_n)T_t)|r_t = r_1, r_0\right)$ is given by

$$E^{\mathrm{P}^h}\left(\exp(-(\rho-\lambda_n)T_t)|r_t = r_1, r_0\right) = \exp(-(\rho-\lambda_n)f(t))L_\gamma(-(\rho-\lambda_n),t)q_r(\rho-\lambda_n,t,r_0,r_1).$$

□

The coefficients $q_r(\rho-\lambda_n,t,r_0,r_1)$ in equation (5.4.10) can be computed as the inverse Fourier transform of the function $G[t,r_0,2\pi i\frac{w}{\delta},\rho-\lambda_n]$ (see remark (5.1.9) and lemma 5.1.7).

The algorithm for computing the price and optimal stopping time of the American option is the following:

(i) Compute the transitional probabilities $p_{\tilde{Y},r}(\Delta,y_0,y_1;r_0,r_1)$.

(ii) At $t = T$ the value function is given by $C^a_\Delta(t,y,r) = g(y)$. The exercise region is the whole two-dimensional lattice $\Lambda = D_r \times D_X$.

(iii) Proceed recursively:

$$C^a_\Delta(t,y,r) = \max\left\{g(y), e^{-\beta\Delta}E^{\mathrm{Q}}\left(C^a_\Delta(t+\Delta,\tilde{Y}_{t+\Delta},r_{t+\Delta})|\tilde{Y}_t = y, r_t = r\right)\right\}, \qquad (5.4.12)$$

where the expectation in equation (5.4.12) is computed as

$$E^{\mathrm{Q}}\left(C^a_\Delta(t+\Delta,\tilde{Y}_{t+\Delta},r_{t+\Delta})|\tilde{Y}_t = y, r_t = r\right) = \sum_{y_1,r_1} p_{\tilde{Y},r}(\Delta,y,y_1;r,r_1)C^a_\Delta(t+\Delta,y_1,r_1).$$

(iv) The continuation region and stopping regions are found as

$$C^T_t = \{(r,y)\in\Lambda : C^a_\Delta(t,y,r) > g(y)\},\quad S^T_t = \Lambda\setminus C^T_t.$$

To conclude this section we will describe some features of this method. There seems to be two major problems when implementing this algorithm. The first problem is of computational nature and it arises when we compute the Hahn polynomials using the recurrence relation (2.2.38). Unfortunately this procedure is not stable; we were able to overcome this difficulty by using Gramm-Schmidt orthogonalization procedure.

The second problem is that the process $Y_t = Y(X_t^h)$ (stochastic transformation of the Hahn process) is a martingale only up to the first time it hits the boundary (that is Y_t stopped at the boundary is a martingale). This problem can be dealt with by choosing parameters of the Hahn process in such a way that the probability of hitting the boundary up to maturity T is small.

The main advantage of our method is that it gives the ability to compute node-to-node transition probabilities with almost no computational error; the only errors appear in computation of the discounted transition probabilities for the Meixner process r_t, but these decrease exponentially in k_2 (see remark 5.1.9). It is important to note that the transition probabilities can be computed for arbitrary time interval Δ, which makes our model very useful for pricing financial derivatives with only a finite number of times at which exercise is possible.

Bibliography

[1] M. Abramovitz and I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical tables*. Wiley and Sons, 1972.

[2] C. Albanese, G. Camplolieti, P. Carr, and A. Lipton. Black-Scholes goes hypergeometric. *Risk*, 14(12):99–103, 2001.

[3] D.S. Bates. Jumps and stochastic volatility: exchange rate processes implicit in deutschemark options. *Review of Financial Studies*, 9:69–108, 1996.

[4] A.N. Borodin and P. Salminen. *Handbook of Brownian Motion: Facts and Formulae*. Birkhauser Verlag, Basel, 1996.

[5] J. Cox and S. Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3:145–166, 1976.

[6] B. Dupire. Pricing with a smile. *Risk*, 7:18–20, 1994.

[7] Stein E. and Stein J. Stock price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies*, 4:183–184, 1991.

[8] I.S. Gradshteyn and I.M. Ryzhik. *Tables of integrals,series and products, 5-th edition*. Academic Press, 1994.

[9] J.J. Gray. Fuchs and the theory of differential equations. *Bull. Amer. Math. Soc. (N.S.),* **10** *(1) (1984), 1-26*, 1984.

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

[11] S. Heston. Invisible parameters in options prices. *Journal of Finance*, 48:993–947, 1993.

[12] J. Hull and A. White. The pricing of options on assets with stochastic volatility. *Journal of Finance*, 42:281–300, 1987.

[13] K. Ito and Jr. H.P. McKean. *Diffusion Processes and their Sample Paths*. Springer-Verlag, Berlin, Heidelberg, and New York, 1965.

[14] R. Koekoek and R. F. Swarttouw. The askey-scheme of orthogonal polynomials and its q-analog. *Delft University of Technology, Faculty of Information Technology and Systems, Dep-t of Technical Mathematics and Informatics, Report no. 98-17*, 1998.

[15] M. Rutkowski M. Musiela. *Martingale methods in Financial modelling*. Springer-Verlag, 1998.

[16] D.B. Madan, P. Carr, and E.C. Chang. The variance gamma process and option pricing. *European Financial Review*, 2:79–105, 1998.

[17] D.B. Madan and F. Milne. Option pricing with vg martingale components. *Mathematical Finance*, 1:39–55, 1991.

[18] D.B. Madan and E. Seneta. The variance-gamma model for share market returns. *Journal of Business*, 63:511–524, 1990.

[19] P. Mandl. *Analytical Treatment of One-dimesional Markov Processes*. Springer-Verlag, Berlin, Heidelberg, and New York, 1968.

[20] O. Mazet. Classification des semi-groupes de diffusion sur r associes a une famille de polynomes orthogonaux, in seminare de probabilites, xxx1, lecture notes in mathematics, vol.1665, p.40-54. 1997.

[21] S. Rady. Option pricing in the presence of natural boundaries and a quadratic diffusion term. *Finance and Stochastics*, 1:331–44, 1997.

[22] W. Schoutens. *Stochastic processes and orthogonal polynomials,Vol. 146 of Lecture Notes in Statistics*. Springer-Verlag, New York, 2000.

[23] A.N. Shiryaev. *Optimal stopping rules*. Springer-Verlag, 1978.

[24] D. Williams. *Diffusions, Markov Processes and Martingales I, II.* Wiley and Sons, 1987.

[25] Wang Zikun and Yang Xiangqun. *Birth and Death processes and Markov Chains*. Springer-Verlag, Berlin, Heidelberg, and New York, 1992.