NTNU
Innovation and Creativity

Numerical Path Integration for Lévy Driven Stochastic Differential Equations

Tore Selland Kleppe

Master of Science in Physics and Mathematics
Submission date: May 2006
Supervisor: Arvid Næss, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem Description

Review the theory for stochastic differential equations driven by Lévy processes. Implement and test the numerical path integration method for such processes.

Assignment given: 2006-01-13
Supervisor: Arvid Næss, MATH

Abstract

Some theory on Lévy processes and stochastic differential equations driven by Lévy processes is reviewed. Inverse Fast Fourier Transform (iFFT) routines are applied to compute the density of the increments of Lévy processes. We look at exact and approximate path integration operators to compute the probability density function of the solution process of a given stochastic differential equation. The numerical path integration method is shown to converge under the transition kernel backward convergence assumption. The numerical path integration method is applied on several examples with non-Brownian driving noises and nonlinearities, and shows satisfactory results. In the case when the noise is of additive type, a general path integration code written for Lévy driving noises specified by the Lévy-Khintchine formula is described. In this code the iFFT routine is an integral part. Some preliminary results on path integration in Fourier space are given.

Preface

This is the master thesis submitted as a conclusion of the Master of Science in Physics and Mathematics (specialization in Industrial Mathematics) degree at the Norwegian University of Science and Technology (NTNU). The thesis work was performed during the spring semester 2006 under supervision of Professor Arvid Næss at the Institute of Mathematical Sciences.

Through the last few years, I have been interested in stochastic dynamics and numerical methods for such problems. The blend of mathematics, statistics and numerics has been intriguing to me, and the work has been most rewarding. This thesis should be viewed upon as the continuation of the work done in the fall semester of 2005. Then I wrote a smaller thesis on the numerical path integration method for Brownian motion driven stochastic differential equations. Section 5.1 and figures 5.11 and 5.12 are taken in slightly reworked forms from the former thesis.

I wish to thank several persons for help and input: Professor Arvid Næss for excellent supervision. Professor Harald Krogstad for rewarding discussions concerning the application of inverse Fast Fourier Transform routines on characteristic functions. Professor Nikolai Ushakov, Ass. Professor Harald Hanche-Olsen, Ass. Professor Espen Robstad Jakobsen and Dr.Ing. student Eirik Mo for rewarding discussions. Finally I wish to thank my beautiful girlfriend Ingerid Nordland for her patience with me working all the time.

Trondheim, June 1. 2006
Tore Selland Kleppe

Contents

Chapter 1

Introduction

The last few decades have seen an increased interest in the field of stochastic analysis. The development of mathematical finance has fueled this development. The increased interest in the field has also reviled some of the weaknesses in the classic stochastic calculus based on Brownian motion. The inability of the classic Black-Scholes model to incorporate heavy tails in the log-returns, jumps and other features made embedded by the use of Brownian motion as driving noise has turned the attention of mathematicians working in mathematical finance to more general Lévy processes (Cont & Tankov 2004), (Schoutens 2003). Also other fields of research have lied their eyes on the tools that a more general stochastic calculus provides. This includes civil engineering (Protter & Talay 1997), physics (Applebaum 2004), hydrology and neurophysiology (Daly & Porporato 2006).

The numerical path integration method is used to approximate the probability density of the solution process of a stochastic differential equation. The method is well-described in the case when the driving noise is Brownian motion (see e.g. Naess (2001), Naess & Moe (2000)) but relatively little have been written when the driving noise is a more general Levy process. In this text we attempt to extend the numerical path integration machinery to a wider class of stochastic differential equations. Due to time-constraints on this work, this has only been done to 1-dimensional models.

The text is lied out as follows. Chapter 2 is a brief exposition of some of the analysis and stochastic analysis tools which we use later in the text. The chapter treats the close connection between the Markov processes described in this text and a class of integro-partial differential operators. Also included in this chapter is the development of some notation which we use when we show convergence of the numerical path integration method.

Chapter 3 is a review of Lévy processes and the stochastic calculus associated with such processes. This will form the theoretical foundation for treating stochastic differential equations driven by Lévy processes.

In chapter 4 we look at methods for numerical inversion of characteristic functions associated with stochastic variables. This will be an integral part of the numerical path integration code for additive noise stochastic differential equations described in chapter 6. Moreover the techniques are used to compute densities from the path integration in Fourier space methods described in appendix A.

Chapter 5 describes the path integration method and the numerical path integration method in some detail. The first part treats the rather abstract notion of path integration, and then describes the path integral in a stochastic differential equation setting. Further we look at methods for approximating the path integration operator on a dig-

ital computer, and show convergence under some assumptions. Finally some issues of implementation are treated.
The final chapter 6 treats examples of numerical path integration for different models. The main scope of this chapter is to show that the path integration method can be implemented for a quite wide family of problems, and gives satisfying results. In many of the examples, we describe the corresponding (integro-)PDE-initial value problem that the numerical path integration method solves approximately.
The reader should have some background in real- and Itô-stochastic analysis. Also basic numerical analysis for ordinary differential equations and basic probability theory is assumed known.

Chapter 2

Mathematical Preliminaries

In this chapter we review some basic results from real and stochastic analysis. The reader is assumed to have some background in measure theory and Itô stochastic analysis. More complete accounts and proofs of the following real analysis results can be found in Stroock (1998), Rudin (1976) or Folland (1999), and the stochastic analysis results in Øksendal (2003), Karatzas & Shreve (1988) or Gihman & Skorohod (1972).

2.1 Some Real and Stochastic Analysis

2.1.1 Some Notation

Unless otherwise noted, we use standard notation for $\mathbb{R}^n$, that is, a vector $x \in \mathbb{R}^n$ is written as $x = [x_1, x_2, \ldots, x_n]$. We write $\mathbb{R}^+ = \mathbb{R} \setminus (\infty, 0]$. The min and max operators $f \vee g = \max(f, g)$, $f \wedge g = \min(f, g)$ are defined pointwise for functions f, g mapping to $\mathbb{R}$. We use the notation $\mathbf{1}_S$ for the indicator function - that is, for some $S \subseteq \mathbb{R}^n$,

$$\mathbf{1}_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{elsewhere.} \end{cases} \tag{2.1}$$

The support of a function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\text{supp}(f) = \{x \in \mathbb{R}^n; f(x) \neq 0\}. \tag{2.2}$$

2.1.2 Basic Real Analysis

We denote the σ-algebra consisting of all open sets on $\mathbb{R}^n$ the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$. Moreover, let $\mu(\mathrm{d}x)$ denote the Lebesgue measure on $\mathcal{B}(\mathbb{R}^n)$. We often use the shorthand notation $\mu(\mathrm{d}x) = \mathrm{d}x$ when there is no room for confusion. An even shorter notation is $\int f = \int_{\mathbb{R}} f(x)\mu(\mathrm{d}x)$. The triplet $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ is called the basic measure space on $\mathbb{R}^n$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be measurable with respect to the basic measure space if $f^{-1}(S) \in \mathcal{B}(\mathbb{R}^n)\ \forall\ S \in \mathcal{B}(\mathbb{R})$.

Definition 2.1: *A measurable function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be of class $L^p(\mathbb{R}^n)$ if*

$$\|f\|_p = \left[\int_{\mathbb{R}^n} |f|^p d\mu\right]^{(1/p)} < \infty \tag{2.3}$$

for $p \in [1, \infty)$ and of class $L^\infty(\mathbb{R})$ if

$$\|f\|_\infty = \inf\{C \geq 0;\ \mu(\{|f| > C\}) = 0\} < \infty. \tag{2.4}$$

The families of functions $L^p(\mathbb{R})$, $p \in [1, \infty]$ are normed spaces where the functionals $\|\cdot\|_p$ are the norms.

The integral (2.3) is taken in the Lebesgue-sense. The space $L^2(\mathbb{R}^n)$ have some special properties:

Theorem 2.1: *$L^2(\mathbb{R}^n)$ is an inner product space with inner product*

$$\langle f, g\rangle = \int_{\mathbb{R}^n} f g d\mu. \tag{2.5}$$

for $f, g \in L^2(\mathbb{R}^n)$.

Theorem 2.2: *For any inner product space we have the Cauchy-Schwartz inequality; in the special case of $L^2(\mathbb{R}^n)$ the Cauchy-Schwartz inequality can be written as*

$$|\langle f, g\rangle| \leq \|f\|_2 \|g\|_2 \ \forall\ f, g \in L^2(\mathbb{R}^n) \tag{2.6}$$

We introduce some more notation:

Definition 2.2: *Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a bivariate function. The univariate functions $\|\cdot\|_{x,p}$, $\|\cdot\|_{y,p}$ are defined as*

$$y \to \|f\|_{x,p} = \left[\int_{\mathbb{R}} |f(y, x)|^p \mu(dx)\right]^{(1/p)} \tag{2.7}$$

$$x \to \|f\|_{y,p} = \left[\int_{\mathbb{R}} |f(y, x)|^p \mu(dy)\right]^{(1/p)}. \tag{2.8}$$

and obvious analogs apply for L^∞.

A short hand notation for the differentiation operator will make the notation more compact: :

Definition 2.3: *We write for a univariate function $f : \mathbb{R} \to \mathbb{R}$ where the i-th derivative exist:*

$$D^i(f)(x) = \frac{d^i}{dx^i} f(x) \tag{2.9}$$

Some important classes of functions are the C^α spaces:

Definition 2.4: *A function $f : \mathbb{R} \to \mathbb{R}$ is said to be of class $C^\alpha(\mathbb{R})$ for $\alpha \in \mathbb{N} \cup \{0\}$ if it exists and is continuous, and $D^i f$ is continuous and bounded for $i \leq \alpha$.*

Now some basic stochastic analysis:

2.1.3 Basic Stochastic Analysis

We define in the usual manner the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathcal{P})$ such that $\mathcal{P}(\Omega) = 1$ and $\mathcal{F}_t$ is a σ-algebra such that $\mathcal{F}_t \subseteq \mathcal{F}\ \forall\ t \in [0, \infty)$. Moreover the filtration has the increasing property $\mathcal{F}_s \subseteq \mathcal{F}_t\ \forall\ 0 \leq s \leq t$. A function $X : (\Omega, \mathcal{F}) \to \mathbb{R}^n$ is said to be a stochastic variable if it is measurable with respect to $(\Omega, \mathcal{F})$. A family of stochastic variables $\{X_t\}_t$ is called a stochastic process if it is indexed by either $t \in [0, \infty)$ or $t \in (\mathbb{N} \cup \{0\})$. A stochastic process is said to be adapted to the filtration $\mathcal{F}_t$ if X_t is measurable with respect to $\mathcal{F}_t$ for all t.

Definition 2.5: *A stochastic variable X is said to have a probability density if there exist a function d_X such that for all $S \in \mathcal{B}(\mathbb{R}^n)$ we have that*

$$\mathcal{P}(\{\omega \in \Omega; X(\omega) \in S\}) = \int_S d_X(x)\mu(dx) \tag{2.10}$$

We sometimes denote the density the law of X and use the notation $\mathcal{L}[X] = d_X$.

Some important densities are:

Definition 2.6: *A random variable $X : (\Omega, \mathcal{F}) \to \mathbb{R}$ is said to be Gaussian or Normal if there exist $\mu \in \mathbb{R},\ \sigma \in \mathbb{R}^+$ such that*

$$\mathcal{L}[X](x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \tag{2.11}$$

We use the notation $\mathcal{L}[X] = N(\mu, \sigma^2)$ if X is Gaussian.

Definition 2.7: *A random variable $X : (\Omega, \mathcal{F}) \to \mathbb{R}^+ \cup \{0\}$ has an exponential density if*

$$\mathcal{L}[X](x) = \lambda \exp(-x\lambda) \tag{2.12}$$

where $\lambda \in \mathbb{R}^+$.

We use the notation $\mathcal{L}[X] = EXP(\lambda)$ if X is exponential.

Definition 2.8: *A random variable $X : (\Omega, \mathcal{F}) \to S,\ S \subset \mathbb{R},\ \mu(S) > 0$ has a S-uniform density if*

$$\mathcal{L}[X](x) = \frac{\mathbf{1}_S(x)}{\mu(S)}. \tag{2.13}$$

We use the notation $\mathcal{L}[X] = UNIF(S)$ if X is S-uniform.

Definition 2.9: *The expectation operator $\mathbb{E}$ applied on a stochastic variable X is defined as*

$$\mathbb{E}[X] = \int_\Omega X\, d\mathcal{P}. \tag{2.14}$$

When X has a probability density $d_X(x)$ (2.14) reduces to

$$\mathbb{E}[X] = \int_{\mathbb{R}^n} x\, d_X(x) dx. \tag{2.15}$$

The expectation operator of a mapped stochastic variable $f(X),\ f \in L^\infty$ is given as

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}^n} f(x)\, d_X(x) dx. \tag{2.16}$$

Throughout this text, the vast majority of the stochastic processes to be studied are Markov Processes:

Definition 2.10: *If $t \in (\mathbb{N} \cup 0)$, and for each finite collection of times $t_1 < t_2 < \cdots < t_m < t_{m+1} \in (\mathbb{N} \cup 0)$ and corresponding spatial points $x_1, x_2, \ldots, x_m$ we have the equality*

$$\mathcal{P}(X_{t_{m+1}} \in S | X_{t_1} = x_1 \cap X_{t_2} = x_2 \cap \cdots \cap X_{t_m} = x_m) = \mathcal{P}(X_{t_{m+1}} \in S | X_{t_m} = x_m)\ \forall\ S \in \mathcal{B}(\mathbb{R}^n) \tag{2.17}$$

we say that the process X_t is a Markov chain. Completely analogous, if $t \in [0, \infty)$ and for each finite collection of times $t_1 < t_2 < \cdots < t_m < t_{m+1} \in [0, \infty)$ we have the equality (2.17), we say that the process X_t is a Markov process.

Markov processes have some properties which we shall use extensively throughout this text:

Definition 2.11: *Let X_t be a Markov chain or a Markov process. We denote the measure $K(S, x, t', t),\ t' > t$ with property*

$$\mathcal{P}(X_{t'} \in S | X_t = x) = K(S, x, t', t) = \int_S K(dy, x, t', t) \tag{2.18}$$

the transition measure. If it exist, we denote the function $k(y, x, t', t)$ with property

$$\mathcal{P}(X_{t'} \in S | X_t = x) = \int_S k(y, x, t', t) dy \tag{2.19}$$

the transition kernel of X_t.

Definition 2.12: *If the transition kernel is time-invariant, that is*

$$k(y, x, t', t) = k(y, x, t' + h, t + h)\ \forall\ h \in [0, \infty) \tag{2.20}$$

we say that our Markov process (chain) is time-homogeneous and write $k(y, x, t' - t) = k(y, x, t', t)$.

For a time-homogeneous Markov process (chain) we denote the univariate functions $y \mapsto k(y, x, t)$ the forward transition kernel and $x \mapsto k(y, x, t)$ the backward transition kernel. From the above it is clear that the forward transition kernel form a probability density. For time-homogeneous Markov processes (chains) we have the Chapman-Kolmogorov equation, relating the transition kernels for different time steps:

Theorem 2.3 (Chapman-Kolmogorov Equation)**:** *The transition kernel of a time-homogeneous Markov process has the property*

$$k(y, x, t + t') = \int_{\mathbb{R}^n} k(y, z, t)k(z, x, t')dz \ \forall \ t, t' > 0. \tag{2.21}$$

Throughout the text we use the notation B_t for standard Brownian motion and N_t for the Poisson process with rate λ.

2.1.4 Markov Processes and Integro-PDEs

There is close relation between the law of certain Markov Processes and a class of Integro-partial differential equations. We give an informal sketch of the equations usually referred to as the Kolmogorov forward and backward equations, omitting the regularity conditions. The lines of Applebaum (2004) are followed. We define the so-called infinitesimal generator of a time-homogeneous Markov process:

Definition 2.13: *For any function $f \in C_c^2$, i.e. C^2 with compact support, we define the infinitesimal generator A of the time-homogeneous Markov process as*

$$Af(x) = \lim_{t \to 0} \frac{\mathbb{E}[f(X_t^x)] - f(x)}{t} \tag{2.22}$$

where the superscript correspond to the initial condition $X_0^x = x$. If the transition kernel exists we have that

$$Af(x) = \lim_{t \to 0} \frac{\int_{\mathbb{R}^n} f(y)k(y, x, t)dy - f(x)}{t}. \tag{2.23}$$

Assuming that the process is well-behaving enough, we have the following initial value problem

$$\frac{\partial}{\partial t} v(x, t) = Av(x, t), \ x \in \mathbb{R}^n, \ t \in \mathbb{R}^+ \tag{2.24a}$$

$$v(x, 0) = f(x) \tag{2.24b}$$

where $v(x, t) = \mathbb{E}[f(X_t^x)]$. This is the Kolmogorov backward equation. In our case, the operator A is some differential or integro-differential operator.
We can use equation (2.23) to figure out the corresponding equation for the transition kernel k, again omitting the details. Start with the left hand side of (2.24a).

$$\frac{\partial}{\partial t} v(x, t) = \frac{\partial}{\partial t} \int_{\mathbb{R}^n} f(y)k(y, x, t)\mathrm{d}y = \int_{\mathbb{R}^n} f(y) \frac{\partial}{\partial t} k(y, x, t)\mathrm{d}y. \tag{2.25}$$

It can be shown using the conditional expectation operator (Applebaum 2004) that

$$Av(x, t) = \mathbb{E}[Af(X_t^x)] \tag{2.26}$$

hence

$$Av(x,t) = \int_{\mathbb{R}^n} (Af)(y)k(y,x,t)\mathrm{d}y. \tag{2.27}$$

Since we integrate over the same space we have that

$$\int_{\mathbb{R}^n} f(y)\frac{\partial}{\partial t}k(y,x,t) - (Af)(y)k(y,x,t)\mathrm{d}y = 0 \tag{2.28}$$

Assuming that A has a formal adjoint A^* we get the desired relation

$$\int_{\mathbb{R}^n} f(y)[\frac{\partial}{\partial t}k(y,x,t) - A^*k(y,x,t)]\mathrm{d}y = 0. \tag{2.29}$$

Since C_c^2 is dense in L^1, it can be concluded that the transition kernel obeys the Kolmogorov forward equation

$$\frac{\partial}{\partial t}k(y,x,t) - A^*k(y,x,t) = 0. \tag{2.30}$$

This equation is sometimes referred to as the Fokker-Planck equation in physics-literature. We exemplify this with Itô-diffusions:

Example 2.1: Let $X_t \in \mathbb{R}$ be the solution of the Itô-stochastic differential equation.

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t. \tag{2.31}$$

It is relatively easy to see, using the Itô-formula, that (Øksendal 2003)

$$Af(x) = \left[\frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2} + b(x)\frac{\partial}{\partial x}\right] f(x). \tag{2.32}$$

To find the adjoint operator A^*, recall the definition

$$\langle Af, g\rangle = \langle f, A^*g\rangle. \tag{2.33}$$

Here we use the L^2 inner-product, since $f, g \in C_c^2 \Rightarrow f, g \in L^2$ trivially. Integration by parts with compact support yields:

$$\begin{aligned}\langle Af, g\rangle &= \int_{\mathbb{R}} \left[\frac{1}{2}\sigma(x)^2 D^2 f(x) + b(x)Df(x)\right] g(x)\mathrm{d}x \\ &= -\int_{\mathbb{R}} Df(x)D(\frac{1}{2}\sigma(x)^2 g(x)) + f(x)D(b(x)g(x))\mathrm{d}x \\ &= \int_{\mathbb{R}} f(x)\left[D^2(\frac{1}{2}\sigma(x)^2 g(x)) - D(b(x)g(x))\right]\mathrm{d}x = \langle f, A^*g\rangle. \end{aligned} \tag{2.34}$$

Hence the forward operator is given as

$$A^*f(x) = \left[\frac{1}{2}\frac{\partial^2}{\partial x^2}\sigma^2(x) - \frac{\partial}{\partial x}b(x)\right] f(x) \tag{2.35}$$

The forward transition kernel coincides with what is called the basic solution of partial differential equations (Evans 1998). That is, the forward kernel k solves problems on the form

$$\frac{\partial}{\partial t}u(y,t) = A^*u(y,t) \tag{2.36}$$

$$u(y,0) = \delta(y-x) \tag{2.37}$$

where δ denotes the Dirac δ-function. It can be shown that if we have the basic solutions at hand, $p(y,t)$ given as

$$p(y,t) = \int k(y,x,t)f(x)\mathrm{d}x, \tag{2.38}$$

solves the corresponding problem

$$\frac{\partial}{\partial t}p(y,t) = A^*p(y,t) \tag{2.39}$$

$$p(y,0) = f(y),\ f \in L^1. \tag{2.40}$$

Later it will become apparent that f can be taken to be the initial density of a Markov process X_t, i.e. $\mathcal{L}[X_0] = f$ and $\mathcal{L}[X_t^f] = p(y,t)$ where the superscript denotes the initial condition of the process.

2.2 Smooth Densities and Quasi Densities

In this section we state and prove some results that will be handy in the numerical path integration convergence proof. This subsection can be read simultaneous with section 5.6, and is not essential to other parts of this text.

Definition 2.14: *We define $\mathbb{D}$ to be the space of well-behaving probability densities over $\mathbb{R}$. That is, for each $f \in \mathbb{D}$ we have that*

1. $f \geq 0$
2. $\int f = 1$
3. $f \in C^2(\mathbb{R})$

Lemma 2.1: $\mathbb{D} \subset L^1(\mathbb{R})$

Proof.

$$\|f\|_1 = \int |f| = \int f = 1 < \infty \tag{2.41}$$

□

Lemma 2.2: *Let $f \in \mathbb{D}$. Then f is bounded above, i.e. there exist $B_f \in \mathbb{R}^+$ such that $f(x) \leq B_f \, \forall \, x \in \mathbb{R}$.*

Proof. See e.g. Rudin (1976) Theorem 4.15 □

Lemma 2.3: $\mathbb{D} \subset L^2(\mathbb{R})$

Proof. Let $f \in \mathbb{D}$. Using all three properties gives us that $\mu(\{f \geq 1\}) \leq 1$. Moreover, $f \in \mathbb{D}$ implies that f is bounded by $B_f \in \mathbb{R}^+$. This gives us a bound for $\|f\|_2$:

$$\|f\|_2^2 = \int f^2 \leq \mu(\{f \geq 1\})B_f^2 + \|f^2 \mathbf{1}_{\{f<1\}}\|_1 \leq B_f^2 + \|f\|_1 < \infty \tag{2.42}$$

□

Definition 2.15: *For each function $f \in \mathbb{D}$ and $L, R \in \mathbb{R}$, $L < R$ we define the the L, R-truncation (or short the truncation) of f to be $f^t = f\mathbf{1}_{[L,R]}$. We denote the space of L, R-truncations $\mathbb{D}^t([L, R])$.*

Lemma 2.4: *Let $f^t \in \mathbb{D}^t([L, R])$, then we have that*

1. $f^t \geq 0$
2. $\int f^t \leq 1$
3. $f^t \in C^2([L, R])$

The proof is trivial and omitted.

Lemma 2.5: *Let $f \in \mathbb{D}$. For each truncation f^t we then have that $\|f\|_2 \geq \|f^t\|_2$.*

The proof is trivial and omitted.

Lemma 2.6: *The truncations are dense in $\mathbb{D}$ in the L^2-sense. More precisely; for each $f \in \mathbb{D}$ and each $\epsilon_f^t > 0$ there exists $L, R \in \mathbb{R}$ such that*

$$\|f - f^t\|_2 < \epsilon_f^t \tag{2.43}$$

where f^t denotes the L, R-truncation of f.

Proof. Let $f \in \mathbb{D}$. For all $L^* \in \mathbb{R}$ we have trivially that

$$\int_{-\infty}^{L^*} f(x)\mathrm{d}x = 1 - \int_{L^*}^{\infty} f(x)\mathrm{d}x \tag{2.44}$$

Since

$$\lim_{L^* \searrow -\infty} \int_{L^*}^{\infty} f(x)\mathrm{d}x = 1 \tag{2.45}$$

there exist L^* such that for each $0 < \epsilon_{L^*} < 1$ we have that

$$\int_{L^*}^{\infty} f(x)\mathrm{d}x = 1 - \epsilon_{L^*}. \tag{2.46}$$

Hence

$$\int_{-\infty}^{L^*} f(x)\mathrm{d}x = 1 - 1 + \epsilon_{L^*} = \epsilon_{L^*}. \tag{2.47}$$

Completely analogous arguments lead to the converse relation, namely that for each $0 < \epsilon_{R^*} < 1$ there exists $R^* \in \mathbb{R}$ such that

$$\int_{R^*}^{\infty} f(x)\mathrm{d}x = \epsilon_{R^*}. \tag{2.48}$$

To show density in the L^2-sense define

$$S_L = \{a; f(x) < 1 \,\forall\, x \in [-\infty, a)\} \tag{2.49}$$
$$S_R = \{b; f(x) < 1 \,\forall\, x \in (b, \infty]\}. \tag{2.50}$$

It is clear that $S_L, S_R \neq \emptyset$ since $\lim_{|x|\to\infty} f(x) = 0$, and that S_L, S_R are simply connected. Set

$$L = L^* \wedge \sup S_L \tag{2.51}$$
$$R = R^* \vee \inf S_R. \tag{2.52}$$

Then, if we choose L, R as our truncation limits we have that

$$\|f - f^t\|_2^2 = \int_{\infty}^{L} f(x)^2 \mathrm{d}x + \int_{R}^{\infty} f(x)^2 \mathrm{d}x. \leq \epsilon_{L^*} + \epsilon_{R^*} \tag{2.53}$$

Since we are free to choose $\epsilon_{L^*}, \epsilon_{R^*}$, we have shown existence of L, R such that for each $\epsilon^t > 0$

$$\|f - f^t\|_2 \leq \sqrt{\epsilon_{L^*} + \epsilon_{R^*}} < \epsilon^t \tag{2.54}$$

□

Definition 2.16: *Let $f \in \mathbb{D}$. We call the function f^p a L^2-perturbation (or short a perturbation) if there exist an ϵ^p such that*

$$\|f - f^p\|_2 \leq \epsilon^p \tag{2.55}$$

and the following is fulfilled:

1. $f^p \geq 0$
2. $f^p \in C^2(\mathbb{R})$.

We denote the space of L^2-perturbations $\mathbb{D}^p$.

Definition 2.17: *Let $f^p \in \mathbb{D}^p$, then we define the L, R-truncation of f^p in the obvious manner. The space of L, R-truncated perturbed quasi probability densities is denoted $\mathbb{D}^{pt}$.*

Chapter 3

Lévy Processes and Stochastic Differential Equations

This chapter reviews some important properties of the so-called Lévy processes, and stochastic differential equations driven by such processes. We follow closely the definitions in Protter (2004) in this chapter.

3.1 Lévy Processes

3.1.1 Definition

We first define a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathcal{P})$ in the usual manner. On this space we define the Lévy processes:

Definition 3.1: *An adapted process $X = \{X_t\}_{t\geq 0}$ with $X_0 = 0$ a.s. is a Lévy process if*

- *X has increments independent of the past, that is, $X_t - X_s$ is independent of $\mathcal{F}_s, 0 \leq s < t < \infty$.*
- *X has stationary increments, that is, $X_t - X_s$ has the same distribution as X_{t-s}.*
- *X_t is continuous in probability, that is $\lim_{t\to s} X_t = X_s$, where the limit is taken in probability.*

There is a one-to-one correspondence between Lévy processes and a class of probability distributions, namely the infinitely divisible distributions. To see this, consider the Fourier transform of X_t:

$$\phi_{X_t}(u) = \mathbb{E}[\exp(iuX_t)] \tag{3.1}$$

also known as the characteristic function of the random variable X_t. From definition 3.1 it is easy to see that $\phi_0(u) = 1$ and that $\phi_{s+t}(u) = \phi_s(u)\phi_t(u)$ (independent increments). Due to the independent increment property of Lévy processes, it is possible to write the characteristic function of X_t as a finite or infinite product of characteristic functions. It is also easy to see that this implies that an infinitely divisible distribution can be written as a finite or infinite convolution of transition kernels. Sato (1999) shows the one-to-one correspondence stringently.

3.1.2 The Lévy-Khinchine Representation

One important property of the Lévy processes is that any Lévy process can be represented by a triplet consisting of a matrix, a vector and a measure. More precisely (see e.g. Cont & Tankov (2004) Theorem 3.1)

Theorem 3.1: *Let $X = \{X_t\}_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$ with Lévy triplet (A, γ, ν), then the characteristic function is given by:*

$$\phi_{X_t}(u) = \exp\left(t\left[i\gamma^\top u - \frac{1}{2}u^\top Au + \int_{\mathbb{R}^d} \exp(iu^\top x) - 1 - iu^\top x \mathbf{1}_{\{|x|<1\}}(x)\nu(dx)\right]\right) \quad (3.2)$$

where $A \in \mathbb{R}^{d\times d}$ and positive definite, $\gamma \in \mathbb{R}^d$ and ν is a measure on $\mathbb{R}^d$ satisfying

$$\int_{\mathbb{R}^d\setminus\{0\}} (1 \wedge |x|^2)\nu(dx) < \infty \quad (3.3)$$

Many important properties of a Lévy process X_t can be found by studying the Lévy triplet and the Lévy-Khinchine representation. It is easy to see that the two first terms in the exponent corresponds to constant drift and Brownian motion. The last term determines the jumps of the process. The measure ν, called the Lévy measure, has a simple interpretation in the one dimensional case if $\nu(\mathbb{R}) < \infty$. Then jumps of size in $S \subset \mathbb{R}$ occur according to a Poisson process with intensity parameter $\nu(S)$ (Schoutens 2003). Similar interpretations in more dimensions apply. To explore the Lévy triplet further, consider the following examples in one dimension i.e. $d = 1$:

Example 3.1: Let the Lévy triplet for $X = \{X_t\}_{t\geq 0}$ be given as $(\sigma, \beta, 0)$, that is $\nu(S) = 0$ for all Borel sets $S \subset \mathbb{R}$ and $\sigma, \beta \in \mathbb{R}$, then $X_t = X_0 + \sigma B_t + \beta t$ where B_t is standard Browian motion. This process is called Browian motion with drift.

Example 3.2: The Poisson process is another Lévy processes, being a pure jump process with jumps of fixed size 1 and intensity λ. The Lévy triplet for this process is given as $(0, 0, \lambda\delta_1(\mathrm{d}x))$. δ_1 denotes the Dirac δ-measure on 1.

Example 3.3: An important class of Lévy processes which we use as building blocks for general Lévy processes is the compound Poisson processes. We define them in the following way. Let N_t be an ordinary Poisson process with intensity parameter λ. That is the stochastic process taking values in the non-negative integers with the discrete probability measure

$$\mathcal{P}[N_t = j|N_0 = 0] = \exp(-\lambda t)\frac{(\lambda t)^j}{j!},\ j = 0, 1, \ldots. \quad (3.4)$$

Moreover let $Z_j,\ j = 1, 2, \ldots$ be a sequence of independent random variables with identical law and characteristic function ϕ_Z. We define the compound Poisson Y_t process as

$$Y_t = \sum_{j=1}^{N_t} Z_j \quad (3.5)$$

where the empty sum is defined to be 0.
To show that the compound Poisson process is a Lévy process, we compute the characteristic function

$$\phi_{Y_t}(u) = \mathbb{E}[\exp(iuY_t)] = \mathbb{E}\left[\exp\left(iu\sum_{j=1}^{N_t} Z_j\right)\right] \tag{3.6}$$

$$= \sum_{k=0}^{\infty} \mathcal{P}(N_t = k)\mathbb{E}\left[\exp\left(iu\sum_{j=1}^{k} Z_j\right)\right] \tag{3.7}$$

$$= \sum_{k=0}^{\infty} \exp(-\lambda t)\frac{(\lambda t)^k}{k!}\{\mathbb{E}[\exp(iuZ)]\}^k \tag{3.8}$$

$$= \exp(-\lambda t)\exp\left(\lambda t\mathbb{E}[\exp(iuZ)]\right) \tag{3.9}$$

$$= \exp\left(\lambda t\mathbb{E}[\exp(iuZ) - 1]\right). \tag{3.10}$$

Let $M(\mathrm{d}x) = \mathcal{P}[Z \in \mathrm{d}x]$, then

$$\phi_{Y_t}(u) = \exp\left(\lambda t\int_{\mathbb{R}} (\exp(iux) - 1)M(\mathrm{d}x)\right). \tag{3.11}$$

Setting $\nu(\mathrm{d}x) = \lambda M(\mathrm{d}x)$ we obtain the Lévy-Khintchine representation of the process:

$$\begin{aligned}\phi_{Y_t}(u) &= \exp\left(t\int_{\mathbb{R}} (\exp(iux) - 1)\nu(\mathrm{d}x)\right) \\ &= \exp\left(t\left[iu\int_{[-1,1]} x\nu(\mathrm{d}x) + \int_{\mathbb{R}} (\exp(iux) - 1 - iux\mathbf{1}_{[-1,1]}(x))\nu(\mathrm{d}x)\right]\right).\end{aligned} \tag{3.12}$$

Assume that there exist some $d_Z = \mathcal{L}(Z)$. Then we have that $\nu(\mathrm{d}x) = (d_Z(x)\mathrm{d}x)/\lambda$ and the Lévy triplet can be written as $(0, \int_{[-1,1]} x(d_Z(x)\mathrm{d}x)/\lambda, (d_Z(x)\mathrm{d}x)/\lambda)$.

3.1.3 Properties of the Paths of Lévy Processes

First we define an important class of stochastic processes:

Definition 3.2: *A process X_t is said to be cadlag (from French: "Continu à droite, limite à gauche") if it is right continuous with left limits; that is*

$$X_{t-} = \lim_{s \nearrow t} X_s \tag{3.13}$$

and

$$X_{t+} = \lim_{s \searrow t} X_s \tag{3.14}$$

exist, and $X_t = X_{t+}$.

Protter (2004) shows that every Lévy process has a unique modification that is cadlag. In the rest of the text we shall only discuss the cadlag modification of the Lévy process in question.
From the definition above, it is clear that the jump at time s, which we denote $\Delta X_s = X_{s+} - X_{s-} = X_s - X_{s-}$ is well-defined for cadlag processes.
To gain the full understanding of how the Lévy measure works, the following relation is useful (Protter 2004): Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be bounded and vanishing in a neighborhood of 0, then

$$\mathbb{E}\left[\sum_{0<s\leq t} f(\Delta X_s)\right] = t\int_{-\infty}^{\infty} f(x)\nu(\mathrm{d}x). \tag{3.15}$$

The sum on the left hand side is taken over every time s when a jump occurs. Taking $f(x) = \mathbf{1}_S(x)$ where S is some set in $\mathbb{R}^d$ the relation tells us that the expected sum of jumps $\Delta X_s \in S$ is the integral over S with the Lévy measure.

3.2 Stochastic Differential Equations Driven by Lévy Processes

This section is devoted to develop stochastic calculus for Lévy processes. We follow the lines of Protter (2004), and develop stochastic calculus for a class of stochastic processes called semimartingales. The connection to Lévy processes will become apparent later on. Due to time and space constraints on this work, many details are skipped.

3.2.1 Semimartingales and the Stochastic Integral

We start off with the somewhat involved definition of semimartingales. Semimartingales are roughly speaking the class of stochastic processes where the we can define a stochastic integral in the same manner as the Itô integral. Let us start with the integrand:

Definition 3.3: *A process H is said to be simple predictable if H has a representation*

$$H_t = H_0 \boldsymbol{1}_{\{0\}}(t) + \sum_{i=1}^{n} H_i \boldsymbol{1}_{(T_i, T_{i+1}]}(t) \tag{3.16}$$

where $0 = T_1 \leq \cdots \leq T_{n+1} < \infty$ is a finite collection of stopping times with respect to $\mathcal{F}_t$, H_i is adapted to $\mathcal{F}_{T_i}$ and $|H_i| < \infty$ a.s. for $0 \leq i \leq n$. The family of simple predictable processes is denoted $\boldsymbol{S}$.

We denote $\mathbf{S}$ with uniform convergence in (t, ω) as topology $\mathbf{S}_u$. Moreover let $\mathbf{L}^0$ be the space of finite valued random variables topologized by convergence in probability. Between these two spaces we define the map which will be our integral for simple predictable

processes given a reasonable process X. More precisely we define $I_X(H) : \mathbf{S} \to \mathbf{L}_0$ as

$$I_X(H) = H_0 X_0 + \sum_{i=1}^{n} H_i (X_{T_{i+1}} - X_{T_i}). \tag{3.17}$$

Definition 3.4: *A process X_t is a total semimartingale if X is cadlag, adapted and $I_X(H) : \boldsymbol{S} \to \boldsymbol{L}_0$ is continuous.*

Let X_t be a process and T a stopping time, then the notation X_t^T denotes the process $X_t^T = \{X_{t \wedge T}\}_{t \geq 0}$.

Definition 3.5: *A process is a semimartingale if, for each $t \in [0, \infty)$, X^t is a total semimartingale.*

Hence a semimartingale is a process that gives meaning to the map I_X defined on $\mathbf{S}_u$ as an integral for arbitrary finite integration limits.
Following the usual routine of extending integrals of simple functions to more general spaces, Protter (2004) then shows that $\mathbf{S}$ dense in the space of cadlag processes. More precisely:

Definition 3.6: *Let $\boldsymbol{D}$ and $\boldsymbol{L}$ denote two spaces of adapted processes with cadlag paths.*

On these spaces we define a topology:

Definition 3.7: *A sequence of processes $\{H^n\}_{n \geq 1}$ converges to H uniformly on compacts in probability (UCP) if for each $t > 0$, $\sup_{0 \leq s \leq t} |H_s^n - H_s|$ converges in probability.*

Further Protter (2004) shows that

Theorem 3.2: *The space $\boldsymbol{S}$ is dense in $\boldsymbol{L}$ under UCP topology.*

Finally we define the stochastic integral for a process in $\mathbf{L}$:

Definition 3.8: *Let X be a semimartingale. Then the continuous linear mapping $I_{X^t} : \boldsymbol{L}_{UCP} \to \boldsymbol{D}_{UCP}$ obtained as the extension of $I_{X^t}(H) : \boldsymbol{S} \to \boldsymbol{D}$ is called the stochastic integral.*

Remark 3.1. From (3.17) it is clear that the Itô integral is one instance of the stochastic integral, since the integrand is evaluated in the left endpoint, and obviously X_t is a Brownian motion.

Now that we have the stochastic integral we are ready to develop the notion of stochastic differential equation driven by semimartingales.

3.2.2 Stochastic Differential Equations Driven by Semimartingales

Given the results in the previous subsection, we are ready to give meaning to the equation

$$\mathrm{d}X_t = f(X_{t-})\mathrm{d}Z_t \tag{3.18}$$

where Z_t is a m-vector semimartingale and $Z_0 = 0$. The notation is just the shorthand notation of the integral equation:

$$X_t^i = x^i + \sum_{\alpha=1}^{m} \int_0^t f_\alpha^i(X_{s-})\mathrm{d}Z_s^\alpha \tag{3.19}$$

where $i = 1, \ldots, d$, X_t^i denotes the i-th component of the vector process X and Z_s^α denotes the α-th component of the process Z at time s. The coefficient functions $f_\alpha^i : \mathbb{R}^d \to \mathbb{R}$ are given and we denote $f(x)$ the $d \times m$ matrix function $(f_\alpha^i(x))$ admitting the notation

$$X_t = x + \int_0^t f(X_{s-})\mathrm{d}Z \tag{3.20}$$

which is equivalent to (3.18).
As for ordinary differential equations, we have a theorem that ensures existence and uniqueness of solutions of (3.18). First we need the notion of locally Lipschitz functions:

Definition 3.9: *A function $f : \mathbb{R}^d \to \mathbb{R}$ is said to be locally Lipschitz if there exist constants C_K only dependent on K such that*

$$|x - y| < K \Rightarrow |f(x) - f(y)| < C_K \tag{3.21}$$

for every $x, y \in \mathbb{R}^d$ and $|\cdot|$ being the Euclidean norm.

It is rather obvious that functions of class C^1, that is continuous functions with continuous partial derivatives, are locally Lipschitz. We are finally ready to state the main theorem (Theorem V.38 in Protter (2004)):

Theorem 3.3: *Let Z and f be as above with f locally Lipschitz. Then there exists a function $\zeta(x, \omega) : \mathbb{R}^d \times \Omega \to [0, \infty]$ such that each x, $\zeta(x, \cdot)$ is a stopping time, and there exists a unique solution of*

$$X_t = x + \int_0^t f(X_{s-})\,dZ_s \tag{3.22}$$

up to $\zeta(x, \cdot)$ with $\limsup_{t \to \zeta(x,\cdot)} \|X_t\| = \infty$ a.s. on $\{\zeta < \infty\}$.

Hence stochastic differential equations with locally Lipschitz coefficient functions have unique solutions up to explosion times $T(\omega) = \zeta(x, \omega)$.
It is also worth noticing that if the coefficient functions are taken to be globally Lipschitz, the solutions exists and are unique for all times $t \in [0, \infty)$. This is in accordance with the classical result for Itô stochastic differential equations given in e.g. Øksendal (2003).

3.2.3 Stochastic Differential Equations Driven by Lévy Processes

Finally we are ready to address stochastic differential equations driven by Lévy processes. The following results can be found in Protter (2004):

Theorem 3.4: *A Lévy process is a semimartingale.*

The deterministic function $g(t) = t$ is a semimartingale since it is the Lévy process with Lévy triplet $(0, 1, 0)$. Hence any stochastic differential equation on the form

$$\mathrm{d}X_t = f_1(X_{t-})\mathrm{d}t + f_2(X_{t-})\mathrm{d}L_t \tag{3.23}$$

is well-defined for a (m-vector) Lévy process L_t, and we can apply the results above concerning stochastic differential equations driven by semimartingales.
One important property of (3.23) is that the solution is a strong Markov process. Moreover it is in fact true that if the solution of a general stochastic differential equation (3.18) is a strong Markov process, the driving noise Z is a Lévy process (Protter 2004).

3.2.4 The Infinitesimal Generator

As noted in Protter & Talay (1997), a motivation for the analysis of equations on the form (3.22) is that it can be used to solve the Kolmogorov Backward equation. A nice expression for the generator A of the process X_t solving (3.22) is at hand:

Theorem 3.5 (Protter (2004) Exercise V-8)**:** *Let L_t be a Lévy process with Lévy measure ν. Moreover let L_t have the decomposition $L_t = bt + cB_t + M_t$ where M_t is a pure jump martingale. Then the generator of the process X_t solving (3.22) is given as*

$$\begin{aligned} Ag(x) = \nabla g(x) f(x) b + \frac{1}{2} \sum_{i,j=1}^{d} \left(\frac{\partial^2 g}{\partial x_i \partial x_j}(x) \right) (f(x) c f(x)^\top)^{ij} \\ + \int \nu(dy)(g(x + f(x)y) - g(x) - \nabla g(x) f(x)) \end{aligned} \tag{3.24}$$

where $g \in C_c^\infty$ and $\nabla g(x)$ is a row vector.

No general adjoint operator is to our knowledge given in closed form. However we shall see later that it is often easy to find when a specific equation is given.

3.2.5 A Solvable Stochastic Differential Equation

Example 3.4: An important class of stochastic differential equations with applications in mathematical finance are on form (Cont & Tankov 2004):

$$X_t = 1 + \int_0^t X_s \mathrm{d}L_s. \tag{3.25}$$

Here Z is a 1-dimensional Lévy-process with Lévy triple (σ^2, γ, ν). Its solution is given by the so-called stochastic exponential for rather obvious reasons:

$$X_t = \mathcal{E}(L_t) = \exp[L_t - \sigma^2 t/2] \prod_{0<s\leq t} \{(1 - \Delta L_s) \exp[-\Delta L_s]\} \tag{3.26}$$

Notice that for a continuous process, that is with $\nu = 0$, this equation reduces to the ordinary solution of the Black Scholes Equation. In mathematical finance, is often useful to find the stochastic differential equation with solution $Y_t = \exp(L_t)$ where L_t is a given Lévy process. In the one dimensional case, this can be done following Proposition 8.22.2 in (Cont & Tankov 2004):

Proposition 3.1: *Let Y be given as $Y_t = \exp(L_t)$ where L_t is a Lévy process with Lévy triplet (σ^2, γ, ν), then there exists a Lévy process X such that $Y_t = \mathcal{E}(X_t)$. The Lévy triplet of X, $(\sigma_X^2, \gamma_X, \nu_X)$ is given as:*

$$\sigma_X = \sigma \tag{3.27}$$

$$\gamma_X = \gamma + \frac{\sigma^2}{2} + \int_{\mathbb{R}} \{(e^x - 1)\mathbf{1}_{[-1,1]}(e^x - 1) - x\mathbf{1}_{[-1,1]}(x)\}\nu(dx) \tag{3.28}$$

$$\nu_X(S) = \int_{\mathbb{R}} \mathbf{1}_S(1 - e^x)\nu(dx) \text{ for } S \subset \mathbb{R} \tag{3.29}$$

We see that Lévy stochastic calculus is somewhat more complicated than the Itô stochastic calculus. Generally we do not get nice expressions involving the driving noise only as the strong solution.

3.3 Pathwise Numerical Solution of Stochastic Differential Equations

As for ordinary differential equations, only a small class of stochastic differential equations have analytical solutions in the strong sense. To analyze more complicated stochastic differential equations, one has to turn to numerical methods. The majority of the literature on this subject consider a Monte Carlo approach to find weak solutions of stochastic differential equations by approximating a large ensemble of single paths, and use statistical methods to approximate functionals of the weak solutions. To exemplify this, let $h \in L^\infty$, and let X_t be the solution of a given stochastic differential equation.

We are often interested in the value of $\mathbb{E}[h(X_T)]$ for some fixed time T. To estimate this we use the ordinary mean estimator:

$$\mathbb{E}[h(X_T)] \approx \frac{1}{N}\sum_{j=1}^{N} h(\bar{X}_T^j) \tag{3.30}$$

for an ensemble of approximate paths $\{\bar{X}_t^j\}_j$.
Finding approximate paths is done by discretizing the time, and simulating the random increments using a pseudorandom number generator on a digital computer.

3.3.1 The Euler Scheme

The most studied scheme for Lévy driven stochastic differential equations is the Euler scheme (see e.g. Protter & Talay (1997), Jacod, Kurtz, Mèlèard & Protter (2005) or Kloeden & Platen (1999) in the Brownian motion case). The scheme generalizes the Euler scheme for ordinary differential equations, by including the random increments characterizing the stochastic differential equation. Consider the (vector) stochastic differential equation given by

$$X_t = x + \int_0^t f(X_{s-})\mathrm{d}L_s,\ t \in [0,T],\ T < \infty \tag{3.31}$$

and let $t_i = i\tau$, $i = 0, \ldots, M$ where $\tau = T/M$. Then the Euler Scheme is given recursively as

$$\bar{X}_{t_0}^j = x \tag{3.32}$$

$$\bar{X}_{t_i}^j = \bar{X}_{t_{i-1}}^j + f(\bar{X}_{t_{i-1}}^j)[L_{t_i}(\omega_j) - L_{t_{i-1}}(\omega_j)]. \tag{3.33}$$

The reader will recognize that in the time homogeneous case, the Lévy increment $[L_{t_i}(\omega_i) - L_{t_{i-1}}(\omega_i)]$ is independent of $\mathcal{F}_{t_{i-1}}$. Using this, it is quite straight forward to simulate single paths with the Euler Scheme, provided that we can simulate the increment. Jacod et al. (2005) also treat the case when we can only simulate the increment approximately.

3.3.2 Weak Convergence of the Euler Scheme

The notion of weak convergence of pathwise schemes is important when we consider the convergence of path integration schemes later on: Several notions of weak convergence exist. First we state that of Kloeden & Platen (1999), which is similar to that of Milstein (1995).

Definition 3.10: *A pathwise scheme $\bar{X}_t$ approximating the solution X_t of a given stochastic differential equation is said to have weak convergence of order δ if there exists $C_h < \infty$ not dependent on τ such that*

$$|\mathbb{E}[h(\bar{X}_T)] - \mathbb{E}[h(X_T)]| < C_h \tau^{\delta} \tag{3.34}$$

for all $h \in C_P^{2(\delta+1)}$. Here the subscript "P" denotes that h has at most polynomial growth.

The main results in (Protter & Talay 1997) gives bounds for

$$|\mathbb{E}[h(\bar{X}_T)] - \mathbb{E}[h(X_T)]| \tag{3.35}$$

when h is specified in advance under constraints on f and h. Though they present several theorems for stochastic differential equations under even stronger assumption, we shall only state their most general theorem.
We consider the stochastic differential equation (3.31) where Z_t has Lévy triplet (A, γ, ν). Define for $K > 0$, $m > 0$ and $p \in \mathbb{N} \setminus \{0\}$

$$\begin{aligned}\rho_p(m) = 1 + \|\gamma\|^2 + \|A\|^2 + \int_{-m}^{m} \|z\|^2 \nu(\mathrm{d}z) \\ + \|\gamma\|^p + \|A\|^p + \left(\int_{-m}^{m} \|z\|^2 \nu(\mathrm{d}z)\right)^{p/2} + \int_{-m}^{m} \|z\|^p \nu(\mathrm{d}z).\end{aligned} \tag{3.36}$$

Moreover we define

$$\eta_{K,p}(m) = \exp(K \rho_p(m)) \tag{3.37}$$

and

$$j(m) = \nu(\{x; \|x\| \geq m\}). \tag{3.38}$$

Theorem 3.6: *Suppose that:*

1. $f \in C^4$
2. $h \in C^4$
3. $X_0 \in L^4(\Omega)$ *, i.e. X_0 has finite 4th moment.*

Then there exists a strictly increasing function K depending only on the dimensions of f and the L^∞-norm of the partial derivatives of f and h up to order 4 such that for any discretization step of type T/n, for any integer m

$$|\mathbb{E}[h(\bar{X}_T)] - \mathbb{E}[h(X_T)]| \leq 4\|h\|_\infty (1 - \exp(-j(m)T)) + \frac{\eta_{K(T),8}(m)}{n}. \tag{3.39}$$

More heuristically, Protter & Talay (1997) state that the weak order of convergence is generally lower than one, depending on how the Lévy measure tends to 0 at the tails. Moreover, their results generalize the results in (Kloeden & Platen 1999) which states that the Euler Scheme has weak order $\delta = 1$ when $\nu(S) = 0$ for all $S \subset \mathbb{R}^d$.

3.4 Concluding Remarks

The main objective of this work is to approximate the law of the solution process of a Lévy driven stochastic differential equation. The discussions above form the theoretical basis for this to be meaningful. From now on; all stochastic differential equations under consideration are on the form (3.23) with well-behaving coefficient functions $f_1,\ f_2$.
One of the underlying ideas of the numerical path integration procedure is that we approximate the law (i.e. the transition kernel) of the solution-process of the Euler-scheme. Hence the Euler-scheme will become an integral part of our later discussion.

Chapter 4

Numerical Inversion of the Characteristic Function

This chapter discusses the application of the Fast Fourier Transform to estimate the forward transition kernel for stochastic differential equations driven by Lévy Processes. In this chapter we consider only the 1-dimensional case, but generalizations to vector valued stochastic processes should be straight forward.

4.1 The Fourier Transform

4.1.1 Some Basic Facts

Fourier analysis is a vast field and this account is by no means complete. We recall the definition of the Fourier transform:

Definition 4.1: *Choose $A \in (0, \infty]$ and set $S = [-A, A]$. The functions $f, \hat{f} \in L^1(S)$ form a Fourier pair if*

$$\hat{f}(u) = \int_S \exp(iux) f(x) dx \tag{4.1a}$$

$$f(x) = \frac{1}{2\pi} \int_S \exp(-iux) \hat{f}(u) du. \tag{4.1b}$$

The function $\hat{f}$ is called the Fourier transform of f and conversely f is called the inverse Fourier transform of $\hat{f}$. A short hand notation for the two operators is $\mathscr{F} f = \hat{f}$, $\mathscr{F}^{-1} \hat{f} = f$.

Now we state one of the big theorems of Fourier analysis:

Theorem 4.1: *The Fourier transform is a continuous operator $\mathscr{F} : L^1(\mathbb{R}) \to L^\infty(\mathbb{R})$.*

Proof. See e.g. theorem 8.1 in Champeney (1987) □

From definition 4.1 it is relatively easy to see that a probability density and its characteristic function form a Fourier pair:

Theorem 4.2: *The probability density $f_X(x)$ associated with the stochastic variable X and its characteristic function $\phi_X(u)$ defined as*

$$\phi_X(u) = \mathbb{E}[\exp(iuX)] \tag{4.2}$$

form a Fourier pair. Here $S = supp(f_X)$.

Proof.

$$\phi_X(u) = \mathbb{E}[\exp(iuX)] = \int_S \exp(iuX) f_X(x) \mathrm{d}x = \hat{f}_X(u) \tag{4.3}$$

□

The following result is simple but will become very useful in our later discussions of numerical path integration schemes based on characteristic functions:

Result 4.1: *Take $a \in \mathbb{R}$ and let $\phi(u)$ be the characteristic function associated with the stochastic variable X. Then the characteristic function associated with the stochastic variable aX is given as*

$$\phi_{aX}(u) = \phi_X(au) \tag{4.4}$$

Proof.

$$\phi_{aX}(u) = \mathbb{E}[\exp(iuaX)] = \phi_X(au) \tag{4.5}$$

□

4.1.2 Discrete Fourier Transforms

As in most cases when computers are used to do mathematics, some kind of discretization has to be done. This is also usually the case when the integrals in (4.1) have to be evaluated. Such discretizations are called Discrete Fourier Transforms, and they are no more than weighted trapezoid rules of numerical integration. More precisely, consider $S = [-A, A]$ and discretizise S into N equidistant partitions: $\Delta x = 2A/N$ and $x_j = j\Delta x - A$, $j = 0, \ldots, N-1$. We assume f to be periodic, and that A is chosen such that $f(x_0) = f(x_N)$. Next we define the corresponding discretization of u: $u_k = -\pi N/2A + k\pi/A$, $k = 0, \ldots, N-1$, and we are ready to give the approximation:

$$\hat{f}(u_k) = \int_S \exp(iu_k x) f(x) \mathrm{d}x \approx \Delta x \sum_{j=0}^{N-1} \exp(iu_k x_j) f(x_j) = \hat{F}_k,\ k = 0, \ldots, N-1. \tag{4.6}$$

We use the right hand side as our definition of the discrete Fourier transform and the inverse discrete Fourier transform is defined in a completely analogous way. For computational purposes it is convenient to take N to be a power of 2, since we can use

the so called Fast Fourier Transform (FFT) in this case. See e.g. Gasquet & Witomski (1999) for a more complete account of the FFT. The FFT is an algorithm that computes $\{\hat{F}_k\}_{k=0}^{N-1}$ in only $O(N\log_2(N))$ operations. This is a considerable speed up compared to the $O(N^2)$ operations needed to compute the N sums directly.

4.2 Inverse FFT Applied on the Characteristic Function of a Lévy Process

In this section we describe a method, using Theorem 4.2 and the inverse FFT (iFFT), to compute approximations to the forward transition kernel of increments of a Lévy process.

4.2.1 Conjugate Symmetry

This subsection describes an important property of the Fourier transform of real functions, namely the conjugate symmetry. In this text we only consider probability densities that are real, hence all characteristic functions under consideration are conjugate symmetric. We state this more formally:

Theorem 4.3: *Let $d : \mathbb{R} \to \mathbb{R}$, $d \in L^1(\mathbb{R})$ be a probability density. Then its Fourier transform is conjugate symmetric. That is $\hat{d}(u) = \overline{\hat{d}(-u)}$ where $\bar{\cdot}$ denotes the complex conjugate.*

Proof. The proof is simple; by the Euler formula we have that

$$(\mathscr{F}d)(\omega) = \int_{\mathbb{R}} d(x)\cos(\omega x)\mathrm{d}x + i\int_{\mathbb{R}} d(x)\sin(\omega x)\mathrm{d}x \tag{4.7}$$

$$(\mathscr{F}d)(-\omega) = \int_{\mathbb{R}} d(x)\cos(\omega x)\mathrm{d}x - i\int_{\mathbb{R}} d(x)\sin(\omega x)\mathrm{d}x. \tag{4.8}$$

All the integrals exist by theorem 4.1 and are trivially real, and we have shown the desired relation. □

This relation can be shown the other way around for a large class of Lévy process increments. Recall that the characteristic function of an increment of a Lévy process L_t is given by the Lévy-Khinchine formula in one dimension:

$$\phi_{L_t}(u) = \exp\left(t\left[i\gamma u - \frac{1}{2}a^2u^2 + \int_{\mathbb{R}} \exp(iux) - 1 - iux\mathbf{1}_{\{|x|<1\}}(x)\nu(\mathrm{d}x)\right]\right) \tag{4.9}$$

If we fix the argument at some value $u = \omega$, for some fixed t, and compare $\phi_{L_t}(-\omega)$ and $\phi_{L_t}(\omega)$ we get the following property:

Result 4.2: *Assume that* $\nu(\mathbb{R}) \vee \int_{[-1,1]} x\nu(dx) < \infty$ *and that there exists a function* $d \in L^1(\mathbb{R})$ *such that* $\nu(dx) = d(x)dx$. *Then the Lévy-Khinchine representation is conjugate symmetric, i.e.* $\phi_t(u) = \overline{\phi_t(-u)}$

Proof. Fix some $\omega > 0$ and evaluate $\phi_t(\omega)$:

$$\begin{aligned}\phi_{L_t}(\omega) &= \exp\left(t\left[i\gamma\omega - \frac{1}{2}a^2\omega^2 + \int_{\mathbb{R}} \exp(i\omega x) - 1 - i\omega x \mathbf{1}_{\{|x|<1\}}(x)d(x)\mathrm{d}x\right]\right) \\ &= \underbrace{\exp(-t[\sigma^2\omega^2 + \nu(\mathbb{R})])}_{K_1} \exp(ti\gamma\omega) \exp\left(t\int_{\mathbb{R}} \exp(i\omega x)d(x)\mathrm{d}x\right) \exp\left(-ti\omega \underbrace{\int_{[-1,1]} x\nu(\mathrm{d}x)}_{K_2}\right).\end{aligned} \tag{4.10}$$

The corresponding expression for $-\omega$ becomes:

$$\phi_{L_t}(-\omega) = K_1 \exp(ti\omega(K_2 - \gamma)) \exp\left(t\int_{\mathbb{R}} \exp(i(-\omega)x)d(x)\mathrm{d}x\right) \tag{4.11}$$

It is clear that the integrals in the exponents form Fourier transforms. By theorem 4.3 $c_\omega = \overline{c}_{-\omega}$ when $c_\omega = (\mathscr{F}d)(\omega)$. Set $c_\omega = \alpha + i\beta$ and obviously $c_{-\omega} = \alpha - i\beta$. Inserting this into equations (4.10),(4.11) yields

$$\phi_{L_t}(\omega) = K_1 \exp(\alpha) \exp(i(t\omega\gamma - t\omega K_2 + \beta)) = K_3 \exp(iK_4) \tag{4.12}$$
$$\phi_{L_t}(-\omega) = K_1 \exp(\alpha) \exp(-i(t\omega\gamma - t\omega K_2 + \beta)) = K_3 \exp(i(-K_4)) \tag{4.13}$$

Finally the Euler formula gives us

$$\phi_{L_t}(\omega) = K_3(\cos(K_4) + i\sin(K_4)) \tag{4.14}$$
$$\phi_{L_t}(-\omega) = K_3(\cos(-K_4) + i\sin(-K_4)) \tag{4.15}$$
$$= K_3(\cos(K_4) - i\sin(K_4)) = \overline{\phi_{L_t}(\omega)} \tag{4.16}$$

since $K_1, K_2, K_3, K_4 \in \mathbb{R}$. □

The restrictions on the Lévy measure in result 4.2 might seem very tight. However it can be shown that any Lévy process can be approximated with arbitrarily high accuracy by a finite sum of compound Poisson processes and a Brownian motion with drift (see e.g. Cont & Tankov (2004)), each fulfilling the restrictions.
Having established this property, we review the discrete inverse Fourier transform of the characteristic function.

4.2.2 Inverse Discrete Fourier Transform of a Characteristic Function

The inverse discrete Fourier transform of the characteristic function is given as

$$d_t(x_j) = \Delta u \sum_{k=0}^{N-1} \exp(-iu_k x_j)\phi_t(u_k),\ j = 0, \ldots, N-1 \tag{4.17}$$

where $\Delta u = \pi/A$. Using the results in the previous subsection, we know that $\phi_t(u_k) = \overline{\phi_t(u_{N-k})}$ since $-u_k = u_{N-k}$. Numerically this accounts to only evaluating half the characteristic function values. Moreover, this ensures that $d_t(x_j)$ is real since

$$\begin{aligned}\exp(-iu_k x_j)\phi_t(u_k)+\exp(-iu_{N-k}x_j)\phi_t(u_{N-k}) &= \exp(-iu_k x_j)\phi_t(u_k)+\exp(iu_k x_j)\overline{\phi_t(u_k)} \\ &= \exp(-iu_k x_j)\phi_t(u_k) + \overline{\exp(-iu_k x_j)\phi_t(u_k)} = 2\Re[\exp(-iu_k x_j)\phi_t(u_k)]\end{aligned} \tag{4.18}$$

for $k = 1, \ldots, N/2-1$. For $k = N/2 \Rightarrow u_{N/2} = 0$, the elements are trivially real.
As we see from (4.18) it is possible to simplify and speed up the computation of the discrete transform considerably when using the conjugate symmetry property.

4.3 Error Bounds

Since we are computing approximate distributions, some sort of error analysis should be performed. Hughett (1998) gives us explicit error bounds assuming that the characteristic function and the probability density decay at certain rates in the tails:

Theorem 4.4: *Let d be a probability density function and $\phi(u)$ its characteristic function. Suppose that there exist constants B, C, $\alpha > 1$ and $\beta > 1/2$ such that $|d(x)| \leq B|x|^{-\alpha}$ for all $|x| \geq A$ and $|\phi(u)| \leq C|u/2\pi|^{-\beta}$ for all $|u| \geq \pi N/2A$. Let g be the low pass filtered approximation to d with N on $[-A, A]$. Then the sampling error is bounded by*

$$\frac{1}{\sqrt{2A}}\|d - d_s\|_2 \leq \frac{B}{A^\alpha}\frac{2\alpha - 1}{\alpha - 1} \tag{4.19}$$

and the truncation error by

$$\frac{1}{\sqrt{2A}}\|d_s - g\|_2 \leq \frac{C(2A)^{\beta-1}}{\sqrt{\beta - 1/2}}\left(\frac{2}{N-2}\right)^{\beta-1/2} \tag{4.20}$$

The low pass filtered approximation is the one given by (4.17) when varying the x-grid to get a continuous function. In our case it will generally be sufficient to compute the density on a discrete grid, and the details for obtaining a continuous function is therefore omitted.
It is relatively easy to see that under the assumptions of result 4.2 the characteristic function of a Lévy increment will obey the tail bound of theorem 4.4:

Result 4.3: *Let $\phi_t(u)$ be given as in result 4.2. Then there exists a $C < \infty$ such that $|\phi(u)| \leq C|u/2\pi|^{-\beta} \forall\, |u| \geq \pi N/2A$.*

Proof.

$$|\phi_t(u)| = \exp\left(t(\hat{d}(u) - \nu(\mathbb{R}))\right) \exp(-t\sigma^2 u^2)$$
$$\leq \exp\left(t(\|\hat{d}\|_\infty - \nu(\mathbb{R}))\right) \exp(-t\sigma^2 u^2) = K \exp(-t\sigma^2 u^2) \quad (4.21)$$

The inequality is justified by the fact that the Fourier transform is a linear operator $\mathscr{F} : L^1(\mathbb{R}) \to L^\infty(\mathbb{R})$. To show the existence of C, set $b = \pi N/2A$. We may without loss of generality assume that $b \geq 1$. If we take $C = K \exp(-t\sigma^2 b^2) b^\alpha$ and consider $u \geq b$ then

$$\log\left(\frac{C}{u^\alpha}\right) = \log(K) - t\sigma^2 b^2 + \alpha(\log(b) - \log(u)) \quad (4.22)$$

$$\log(K \exp(-t\sigma^2 u^2)) = \log(K) - t\sigma^2 u^2 \quad (4.23)$$

since u^2 grows faster than $\alpha \log(u)$ for $u > 1$ it is clear that

$$\frac{C}{u^\alpha} \geq K \exp(-t\sigma^2 u^2)) \,\forall u \geq b \quad (4.24)$$

Completely analogous arguments give us the same result for $u \leq -b$. □

4.4 Implementation and Numerical Examples

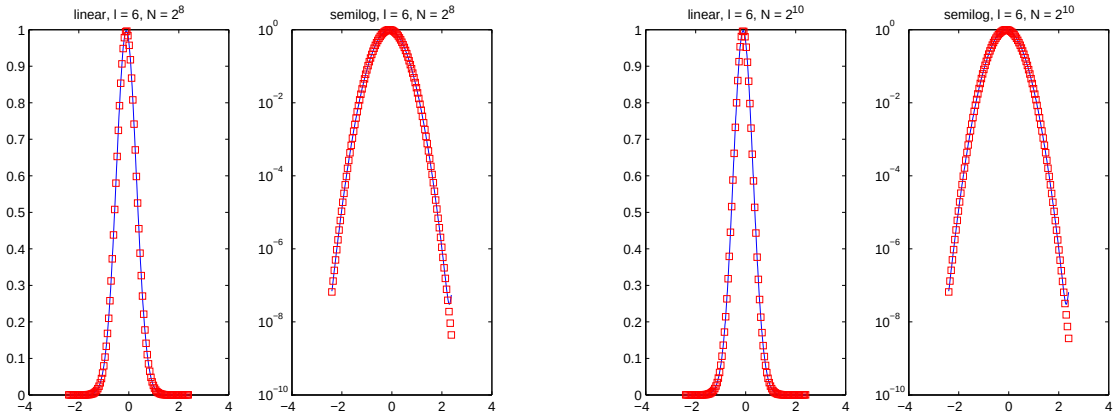
To illustrate the technique outlined above, we consider some numerical examples. The inverse transformations are done using the routine `ZFFTD` in the Sun Performance Library (*Sun Studio 11: Sun Performance Library Reference Manual - ZFFTD* 2005). `ZFFTD` is a routine written for conjugate symmetric input of length N/2+1, and returns output of length N. This accounts to only evaluating $\{\phi_t(u_k)\}_{k=0}^{N/2}$. The code is written in `Fortran 90`.

4.4.1 Finding Appropriate Windows

As in any numerical integration procedure, infinite integration limits are intangible, and we will have to truncate the integration limits to where the integrand is close to numerical zero. In the code used here we choose A to be some multiple of the standard deviation of the stochastic variable. We can approximate the standard deviation by evaluating the characteristic function once. More precisely we have that

$$\mathbb{E}[X] = \mu = -i\frac{\mathrm{d}}{\mathrm{d}u}\phi_X(u)|_{u=0} \quad (4.25)$$

$$\mathbb{E}[(X-\mu)^2] = \sigma^2 = \frac{\mathrm{d}^2}{\mathrm{d}u^2}\phi_X(u)|_{u=0} - \mu^2 \quad (4.26)$$



(a) $l = 6, N = 2^8$ (b) $l = 6, N = 2^{10}$

Figure 4.1: The Gaussian density function computed with $l = 6$ standard deviations on each side. (-) is the iFFT computed density, and (□) is the exact density. Notice that the computed density is periodic.

To estimate the standard deviation σ we use the first and second finite differences:

$$\mu \approx \bar{\mu} = -i\frac{\phi_X(h) - \overline{\phi_X(h)}}{2h} \tag{4.27}$$

$$\sigma^2 \approx \bar{\sigma^2} = \frac{\phi_X(h) - 2 + \overline{\phi_X(h)}}{h^2} - \bar{\mu}^2 \tag{4.28}$$

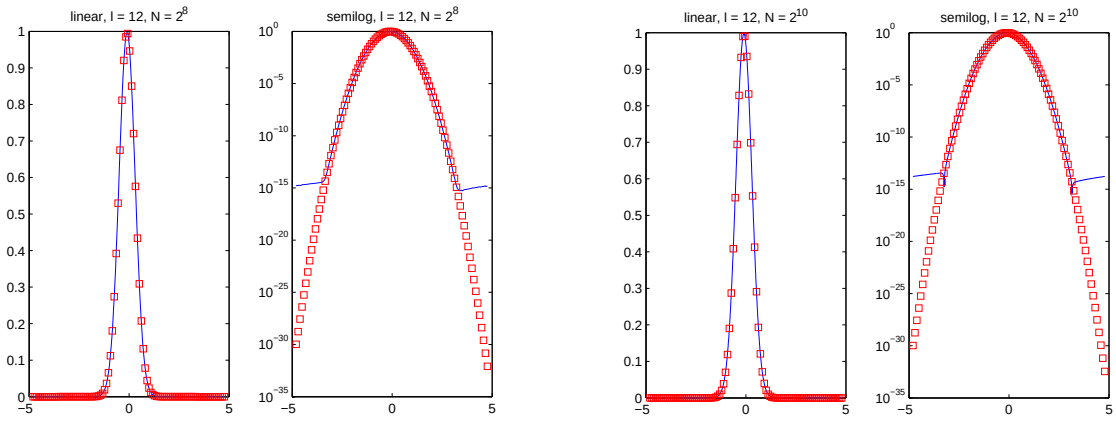
for $1 >> h > 0$. Here we use that $\phi_X(0) = 1$ in the Lévy-Kitchine representation. Hence we call $[-A, A] = [-l\bar{\sigma}, l\bar{\sigma}]$ our window, where l is approximately the number of standard deviations included.

4.4.2 Numerical Examples

Example 4.1: The first example is the Gaussian density. More precisely: Let $\mathcal{L}[X] = N(\mu, \sigma^2)$ and consider the stochastic variable $Y = 2X$ with $\mu = 0.05, \sigma = 0.2$. This is a typical forward transition kernel when working with discretized Itô stochastic differential equations. The characteristic function is trivially given as:

$$\phi_Y(u) = \exp(2i\mu u - \frac{1}{2}(2u\sigma)^2). \tag{4.29}$$

The approximate density given in figures 4.1a - 4.1d for $N = 2^8, 2^{10}$ and $l = 6, 12$. We see that in the case where $l = 6$ the iFFT computed density follows the exact density all the way. When $l = 12$ the iFFT computed density deviates from the exact density at roughly 10^{-14}, which should be accurate enough for most numerical computations. Notice that the iFFT



(c) $l = 12, N = 2^8$ (d) $l = 12, N = 2^{10}$

Figure 4.1: The Gaussian density function computed with $l = 12$ standard deviations on each side. (-) is the iFFT computed density, and (□) is the exact density.

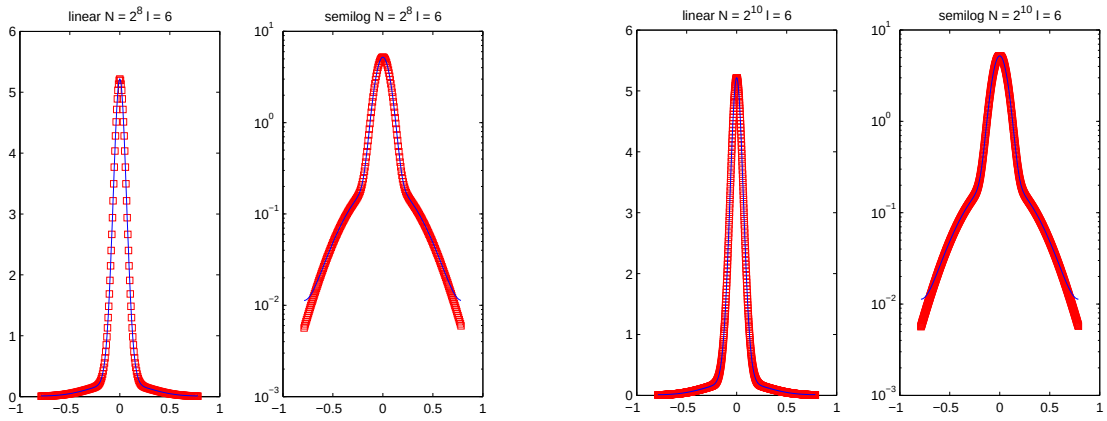
computed density is periodic. This is only a small source of error in case of the Gaussian density since the tails behaves equally at both sides. In the case where the tails have very different behavior, this will be a greater source of error.

Example 4.2: Now consider a jump diffusion increment. In our case $U_\tau = B_\tau + X_\tau$ where B_t is standard Brownian motion, and X_t is a compound Poisson process with $\mathcal{L}[Z_i] = N(0, \sigma_Z^2)$ and rate λ. The variables B_t, N_t and $\{Z_i\}$ are assumed to mutually independent. We derive the exact transition kernel $k(y, x, \tau) = \mathcal{L}[U_\tau | U_0 = x](y)$ in the same manner as in example 3.3. For simplicity we set $x = 0$, then we have that:

$$k(y, 0, \tau) = \exp(\lambda\tau) \sum_{j=0}^{\infty} \frac{(\lambda\tau)^j}{j!\sqrt{2\pi(\tau + j\sigma_Z^2)}} \exp\left(-\frac{y^2}{2(\tau + j\sigma_Z^2)}\right). \tag{4.30}$$

The sum converges fast and approximately 20 terms are sufficient to reach errors less than numerical zero on a typical digital computer. Hence we can use the partial sum as our reference when we look at the iFFT computed densities.
The next step is to derive the characteristic function for such an increment. Recall that the Lévy-measure for a compound Poisson processes is given as $\nu(\mathrm{d}x) = \lambda\mathcal{L}[Z_i](x)\mathrm{d}x$. Hence

$$\phi_{U_\tau}(u) = \exp\left(\tau\left(\lambda\int_{[-1,1]} iux\nu(\mathrm{d}x) - \frac{1}{2}u^2 + \lambda\int_{\mathbb{R}} (\exp(iux) - 1 - iux\mathbf{1}_{[-1,1]}(x))\nu(\mathrm{d}x)\right)\right) \tag{4.31}$$



(a) $l = 6, N = 2^8$ (b) $l = 6, N = 2^{10}$

Figure 4.2: The jump diffusion forward transition kernel computed with $l = 6$ standard deviations on each side. (-) is the iFFT computed density, and (□) is the exact density. Notice that the computed density is periodic.

Since the Lévy-measure is symmetric about the origin we have that

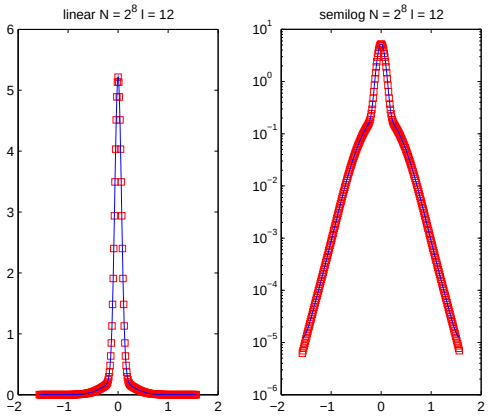
$$\phi_{U_\tau}(u) = \exp\left(\tau\left(-\frac{1}{2}u^2 + \lambda\int_{\mathbb{R}}(\exp(iux) - 1)\frac{1}{\sqrt{2\pi\sigma_Z^2}}\exp\left(-\frac{x^2}{2\sigma_Z^2}\right)\mathrm{d}x\right)\right)$$
$$= \exp\left(-\tau\frac{1}{2}u^2 + \tau\lambda\left(\exp\left(-\frac{1}{2}\sigma_Z^2u^2\right) - 1\right)\right) \quad (4.32)$$

and finally an application of result 4.1 yields

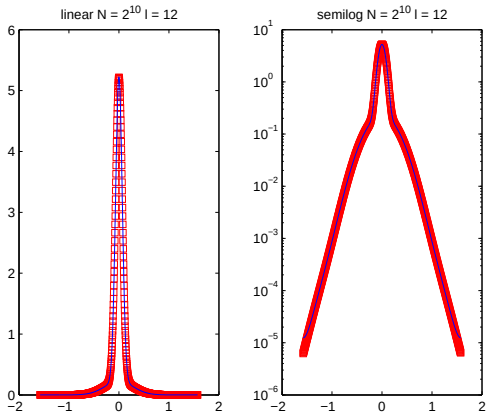
$$\phi_{bU_\tau}(u) = \exp\left(-\tau\frac{1}{2}(bu)^2 + \tau\lambda\left(\exp\left(-\frac{1}{2}\sigma_Z^2(bu)^2\right) - 1\right)\right). \quad (4.33)$$

Hence we have an expression suitable for numerical inversion by iFFT.
In our example we use the parameters $b = 0.3,\ \tau = 0.05,\ \lambda = 3.5,\ \sigma_Z = 0.9$. The iFFT computed densities are plotted in figures 4.2a - 4.2d. We see that the computed density follows the reference density except for the last few points in both tails in all cases.

Example 4.3: The last example is analogous to the previous, but we let $\mathcal{L}[Z_i] = EXP(1/\beta)$. This example is not included to illustrate the accuracy of the iFFT computed densities, but to illustrate what happens when the tails behaves differently. Let us first compute the



(c) $l = 12, N = 2^8$



(d) $l = 12, N = 2^{10}$

Figure 4.2: The jump diffusion forward transition kernel computed with $l = 12$ standard deviations on each side. (-) is the iFFT computed density, and (□) is the exact density.

characteristic function of the increment:

$$\phi_{U_\tau}(u) = \exp\left(\tau\left(-\frac{1}{2}u^2 + \lambda\int_{\mathbb{R}}(\exp(iux)-1)\mathbf{1}_{[0,\infty]}(x)\frac{1}{\beta}\exp\left(-\frac{x}{\beta}\right)\mathrm{d}x\right)\right)$$
$$= \exp\left(-\tau\frac{1}{2}u^2 + \tau\lambda\left(\frac{1}{1-iu\beta}-1\right)\right) \quad (4.34)$$

and

$$\phi_{bU_\tau}(u) = \exp\left(-\tau\frac{1}{2}(bu)^2 + \tau\lambda\left(\frac{1}{1-iu\beta b}-1\right)\right). \quad (4.35)$$

The iFFT computed density is plotted in figure 4.3 with parameters $b = 0.9,\ \lambda = 1.5,\ \beta = 1.1, \tau = 1/20$ and $l = 50$. The computation was done with $N = 2^{10}$. We see from figure 4.3 a rather obvious fact. When the tail behavior differs significantly, the imposed periodicity makes our computed densities inaccurate. In our case, the left tail decays roughly as a Gaussian density ($O(\exp(-x^2))$), whereas the right tail decays roughly as an exponential density ($O(\exp(-x))$). When our density has this property, care should be taken, and we generally need to remove data or move the window such that $d(x_0) \approx d(x_{N-1})$.

4.5 Concluding Remarks

As we have seen in this chapter, in theory we can compute densities with arbitrary accuracy using iFFT. We have tested our method on two relatively simple problems and one hard problem. This has shown us that care should be taken when working with problems that are quantitatively asymmetric. The process of choosing windows

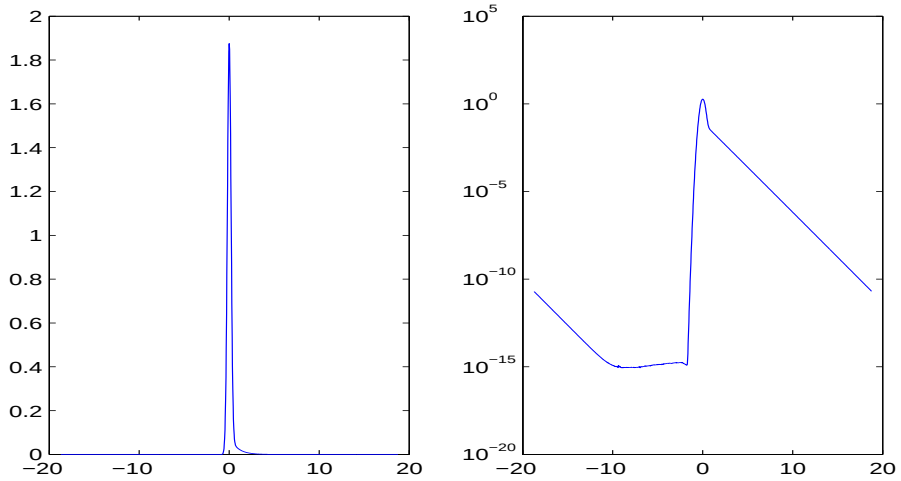


Figure 4.3: The density computed from (4.35) by iFFT. We see that when the tail behaveior differs significantly, the imposed periodicity makes our computed density inacurate.

or discarding data can to some extent be automated by e.g. computing course test densities or comparing gradients in the tails. This has not been implemented in the path integration code which we shall allude to shortly due to time constraints.
The iFFT code used to do to the numerical examples in this chapter has become an integral part of path integration code for additive noise stochastic differential equations described later. Since our Lévy process increments often are specified using the Lévy-Kintchine representation, the alteration of the code to a new model will only consist of changing the characteristic function. In fact, we shall later look at processes where the Lévy process increments have densities equal to those in the examples above.
One important aspect in the numerical examples above is that one generally need a significant part of Brownian motion in the increments to make the computations stable. It is relatively easy to see that pure compound Poisson processes will have a δ-function at the origin for all finite times, making the computations unstable due to scaling problems. This will not cause problems when the Lévy measure is infinite since these processes have smooth forward kernel (Cont & Tankov 2004).

Chapter 5

Path Integration

In this chapter we discuss the exact path integration operator and its time-discrete counterpart in some detail. The understanding of the path integration operator should be that it propagates the probability density of the dynamics specified by some stochastic differential equation trough time. However it is only in very few cases that an analytical expression for the operator is at hand, and generally we have to use some sort of numerical approximation procedure. The last part of the chapter is therefore a description of a numerical path integration procedure. Throughout the chapter we exemplify many of the features discussed with the Langevin equation. Due to time-constraints, the rest of the text is in a 1-dimensional setting.

5.1 Path Integration Heuristics

Path integration in the sense discussed here is a way of computing the time-evolution of some kind of mathematical object given some underlying dynamics. Gas densities and probability distributions are examples of such objects. Path integration was first made famous with Richard P. Feynman's path integral-representation of quantum mechanics (Feng, Wang & Lu 1992). Feynmans principle can, somewhat sloppy, be formulated as that the probability of a particle with position X_t moving from x to A ($A \subset \mathbb{R}^d$), is the integral over all possible paths from x to A with an appropriate probability measure λ:

$$\mathcal{P}(X_T \in A | X_0 = x) = \int_{\Omega_x^A} \lambda(X), \;\; \Omega_x^A = \{X; X_T \in A \,\text{and}\, X_0 = x\}. \tag{5.1}$$

See figure 5.1 for illustration. At first sight this functional integral seems to be a rather intangible object, but we shall see now that there are ways of computing it exactly or by approximation. We start with the Onsager-Machlup Functional method.

5.1.1 Generalized Onsager-Machlup Functional

The Onsager-Machlup Functional is described in Onsager and Machlups paper (Onsager & Machlup 1953) on fluctuations of thermodynamic systems near equilibrium. Onsager and Machlup uses an entropy argument to describe the dynamics of the system subject to additional noise. More mathematically oriented accounts of the Generalized Onsager-Machlup Functional (GOMF) can be found in Wissel (1979) and Risken (1984), where

the underlying dynamics are specified using short time transition probability densities. For simplicity, we consider the GOMF in one space-dimension only in this section. To start our discussion of the GOMF we introduce some notation. Let $a < b$ be two times, and let $a = t_0 < t_1 < \cdots < t_N = b$ be an equidistant partition of $[a, b]$ with step size $\tau = (b - a)/N$. Further; let x_i be the space-variable associated with time i ($x_0 = x$ and $x_N = y$) and let $u(x, a)$ be the initial distribution of the quantity at time a. Fix some $N > 1$ and recall the Chapman-Kolmogorov equation (2.21) for the time-partition. Inserting the short time transition probability

$$k(y, x, \tau) = \frac{1}{\sqrt{2\pi g(x)^2 \tau}} \exp\left(-\frac{[y - x - f(x)\tau]^2}{2g(x)^2\tau}\right) \tag{5.2}$$

corresponding to the stochastic differential equation

$$\mathrm{d}X_t = f(X_t)\mathrm{d}t + g(X_t)\mathrm{d}B_t \tag{5.3}$$

into (2.21) yields:

$$u(y, b) = \int\int\cdots\int \prod_{i=0}^{N-1} \{[2\pi g(x_i)^2\tau]^{-1/2}\mathrm{d}x_i\} \times \exp\left(-\sum_{i=0}^{N-1} \frac{[x_{i+1} - x_i - f(x_i)\tau]^2}{2g(x_i)^2\tau}\right) u(x, a). \tag{5.4}$$

Since the short time transition probabilities are approximations of $O(\tau^2)$ (Risken 1984) we can achieve an explicit formula by letting $N \to \infty$. Before we do so, we use the approximation $x_{i+1} - x_i = \dot{x}(t_i)\tau + O(\tau^2)$, hence the in the limit, the sum in the exponential can be written as:

$$\lim_{N\to\infty} \sum_{i=0}^{N-1} \frac{[x_{i+1} - x_i - f(x_i)\tau]^2}{2g(x_i)^2\tau} = \int_a^b \frac{[\dot{x}(t') - f(x(t'))]^2}{2g(x(t'))^2}\mathrm{d}t'. \tag{5.5}$$

The functional on the right hand side is what is called the Generalized Onsager-Machlup Functional. A common way of writing (5.4) is:

$$u(y, b) = \int_{\Omega_x^{y\mathrm{d}y}} \mathcal{D}_l[x] \exp\left[-\int_a^b \frac{[\dot{x}(t') - f(x(t'))]^2}{2g(x(t'))^2}\mathrm{d}t'\right] u(x, a) \tag{5.6}$$

where $\int_{\Omega_x^{y\mathrm{d}y}} \mathcal{D}$ denotes the infinite integral operator over the function space $\Omega_x^{y\mathrm{d}y}$. The functional in the exponential can be interpreted as a probability measure, assigning a value to all $x \in \Omega_x^{y\mathrm{d}y}$. As noted in Feng et al. (1992) it is generally not possible to compute this functional integral, but some approaches can be followed to yield exact or approximate solutions when more restrictions are imposed on the underlying dynamics. Moreover, Wissel (1979) shows that (5.6) is in fact not uniquely defined when the time

discretization procedure is not specified, but also shows how to work around the mathematical ambiguity of the operator.
One way of achieving an approximate expression for (5.6) , known as the WKB method, is to find the most probable path $x_0 \in \Omega_x^{ydy}$ by a variational argument, and then expand the functional as a functional power series around x_0. Since the first Frechet-derivative (see e.g. Troutman (1996)) is zero, the two first terms in the series simply reduces to a functional over x_0 and the integral operator can generally be tackled. Under certain assumptions on the underlying dynamics, such as linearity in drift and a wellbehaving diffusion matrix, the WKB method turns out to give exact solutions. It is worth noticing that in Onsager & Machlup (1953), the authors apply this method on a linear system to obtain the exact solution.

5.1.2 Generalized Cell Mapping - A Numerical Scheme

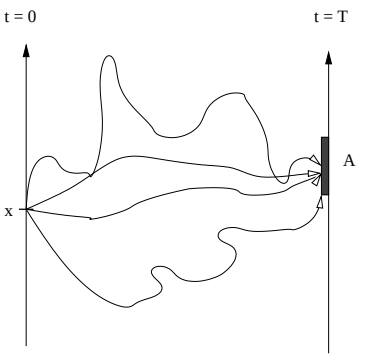


Figure 5.1: Illustration of the path integration principle. We integrate over all possible paths form x to A with an appropriate probability measure.

The generalized Cell Mapping method is a very intuitive numerical scheme for approximating the joint probability density function of Markov processes. The main ideas are to discretize both the state space and the time, and to approximate the original Markov process with a discrete time/space process.
The state space is discretized into a countable number of sets C_i, usually generalized squares, and a probability mass $p_i(n)$ is allocated to set C_i at time step n. The size of the sets is a tradeoff between computation efficiency and desired accuracy. The master equation is nothing but the law of total probability:

$$p_i(n) = \sum_{j=1}^{\infty} P_{ij}^{(n,n-1)} p_j(n-1). \tag{5.7}$$

Notice that in numerically oriented schemes, it is common to work with the joint probability density function of the process (here given as $p_i(n)$) for all time steps. Given (5.7) the process seems straight forward, but the computation of the P_{ij}s can in fact be very hard. Moreover if the dynamics have some time-dependence, the P_{ij}s have to be computed for each time-step. The short time transition probabilities can be approximated by the following:

$$P_{ij}^{(n,n-1)} \approx \int_{C_i} p(x, n\tau | x_j, (n-1)\tau) \mathrm{d}x \tag{5.8}$$

where x_j is the center of C_j and τ is the time step. Different approaches to estimate (5.8) have been used, among them Monte Carlo methods. Sun & Hsu (1990) describes an approach where the short time transition probabilities are Gaussian, with mean and

variance determined by the underlying dynamics. However Sun & Hsu (1990) establish more accurate approximations (but still Gaussian) based on sets of ordinary differential equations describing the two first moments of the transition probabilities, and can therefore take longer time steps compared with similar schemes based on e.g. (5.2)

5.2 Path Integration in a Probabilistic Perspective

As we have seen in the previous section, there is freedom in the way we specify the underlying dynamics of the system we are modeling. In the rest of this text we shall specify the dynamics with time-homogeneous stochastic differential equations of the kind discussed in section 3.2. In this case the measure we apply to each conceivable path from x to S is just the transition kernel k defined for each $S \in \mathcal{B}(\mathbb{R})$:

$$\mathcal{P}[X_{t+\tau} \in S | X_t = x] = \int_S k(y, x, \tau) \mathrm{d}y \ \forall \ t, \tau \geq 0 \tag{5.9}$$

where X_t is the solution of a stochastic differential equation. We start by introducing some notation.

5.2.1 Stochastic Semi-Groups

We define stochastic semi-groups of operators $P_t : L^1(\mathbb{R}) \to L^1(\mathbb{R})$ (Lasota & Mackey 1994):

Definition 5.1: *Let $(X, \mathcal{F}, \mu)$ be a measure space. A Stochastic Semi-Group is a family of operators $P_t : L^1(X) \to L^1(X)$ with the following properties:*

1. $P_t(\alpha f + \beta g) = \alpha P_t f + \beta P_t g, \ f, g \in L^1(X), \ \alpha, \beta \in \mathbb{R}$
2. $P_t f \geq 0$ *if* $f \geq 0$
3. $\int_X P_t f(x) \mu(dx) = \int_X f(x) \mu(dx)$
4. $P_{t+t'} f = P_t(P_{t'} f)$

In addition if

$$\lim_{t \to t'} \| P_t f - P_{t'} f \|_1 = 0 \tag{5.10}$$

the family of operators is called a continuous stochastic semi-group.

This definition admits the path integration operator for a given stochastic differential equation to be a stochastic semi-group. To see this, we define the probabilistic path integration operator (from now on the PI-operator) to be

Definition 5.2: *The PI-operator $PI_t : L^1(\mathbb{R}) \to L^1(\mathbb{R})$ is defined for each initial probability density u_0 as*

$$u_t = PI_t u_0 = \int_{\mathbb{R}} k(y, x, t) u_0(x) dx \tag{5.11}$$

where k is the transition kernel of a given stochastic differential equation.

It is straight forward to check that the PI-operator indeed is a stochastic semi-group:

Result 5.1: *The path integration operator PI_t forms a stochastic semi-group.*

Proof. We check each property of the stochastic semi-group definition:

1. Linearity is obvious since the integral is linear.
2. Also obvious since $k, u_0 \geq 0$.
3. The conservation of probability mass follows from

$$\begin{aligned} \int_{\mathbb{R}} PI_t u_0 = \int_{\mathbb{R}} \int_{\mathbb{R}} k(y, x, t) u_0(x) \mathrm{d}x \mathrm{d}y \\ = \int_{\mathbb{R}} u_0(x) \left(\int_{\mathbb{R}} k(y, x, t) \mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}} u_0 \end{aligned} \tag{5.12}$$

   since

$$\int_{\mathbb{R}} k(y, x, t) \mathrm{d}y = 1 \ \forall \ x \in \mathbb{R} \tag{5.13}$$

4. The last property follows from the Chapman-Kolmogorov Equation.

$$k(y, x, t + t') = \int_{\mathbb{R}} k(y, z, t') k(z, x, t) \mathrm{d}z \tag{5.14}$$

   hence

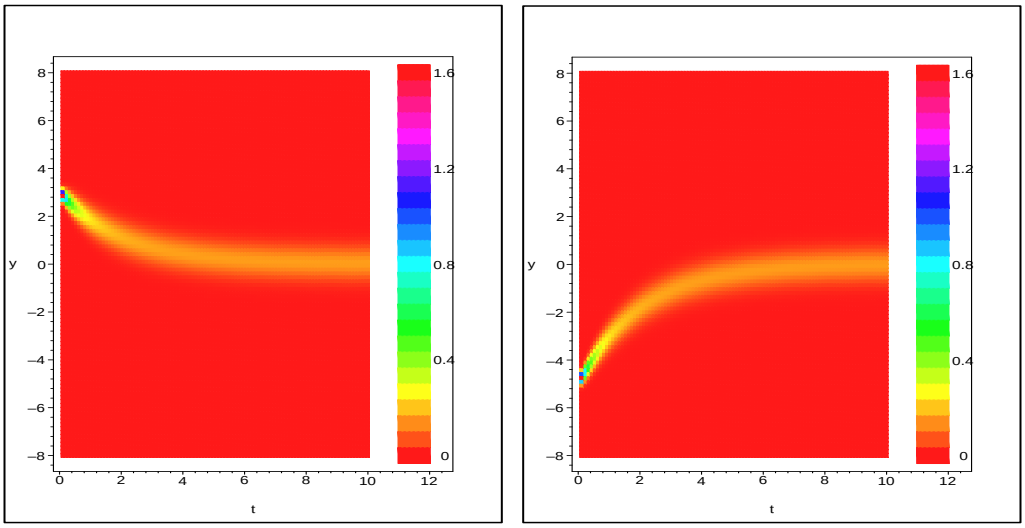
$$\begin{aligned} PI_{t+t'} u_0 = \int_{\mathbb{R}} k(y, x, t + t') u_0(x) \mathrm{d}x = \int_{\mathbb{R}} \int_{\mathbb{R}} k(y, z, t') k(z, x, t) \mathrm{d}z u_0(x) \mathrm{d}x \\ = \int_{\mathbb{R}} k(y, z, t') \left(\int_{\mathbb{R}} k(z, x, t) u_0(x) \mathrm{d}x \right) \mathrm{d}z = PI_{t'}(PI_t u_0) \end{aligned} \tag{5.15}$$

□

To illustrate the principle of path integration we look at an example - the Langevin equation:

Example 5.1: The Langevin equation is the Itô-stochastic differential equation given as

$$\mathsf{d}X_t = -aX_t \mathsf{d}t + b \mathsf{d}B_t, \ X_t \in \mathbb{R}, \ a, b \in \mathbb{R}^+ \tag{5.16}$$



(a) Transition kernel for $x = 3$. (b) Transition kernel for $x = -5$.

Figure 5.2: The transition kernel given in (5.20) for $x = 3, x = -5$ with parameters $a = b = 1/2$.

with initial condition $X_0 = x$. This equation is known to have the strong solution (Kloeden & Platen 1999)

$$X_t = \exp(-at)\left(x + b\int_0^t \exp(as)\mathrm{d}B_s\right) \tag{5.17}$$

Due to the non-stochastic integrand in the Itô-integral it is easy to see that

$$\mathcal{L}\left[x + b\int_0^t \exp(as)\mathrm{d}B_s\right] = N(x, b^2\int_0^t (\exp(as))^2\mathrm{d}s) = N(x, \frac{b^2(\exp(2at)-1)}{2a}) \tag{5.18}$$

hence

$$\mathcal{L}[X_t] = N(x\exp(-at), \frac{b^2}{2a}(1-\exp(-2at))). \tag{5.19}$$

Now from Itô-theory we know that the Itô-integral is a Markov process, hence we can give an explicit expression for the transition kernel k for a general time increment t:

$$k(y,x,t) = \frac{1}{\sqrt{2\pi\frac{b^2}{2a}(1-\exp(-2at))}}\exp\left(-\frac{(y-x\exp(-at))^2}{\frac{b^2}{a}(1-\exp(-2at))}\right). \tag{5.20}$$

The forward transition kernel is plotted in figures 5.2a and 5.2b. To illustrate the path integration principle we use the initial density:

$$u_0(x) = \frac{1}{2\sqrt{2\pi}}\exp\left(-\frac{(x-3)^2}{2}\right) + \frac{1}{2\sqrt{2\pi}}\exp\left(-\frac{(x+5)^2}{2}\right), \tag{5.21}$$

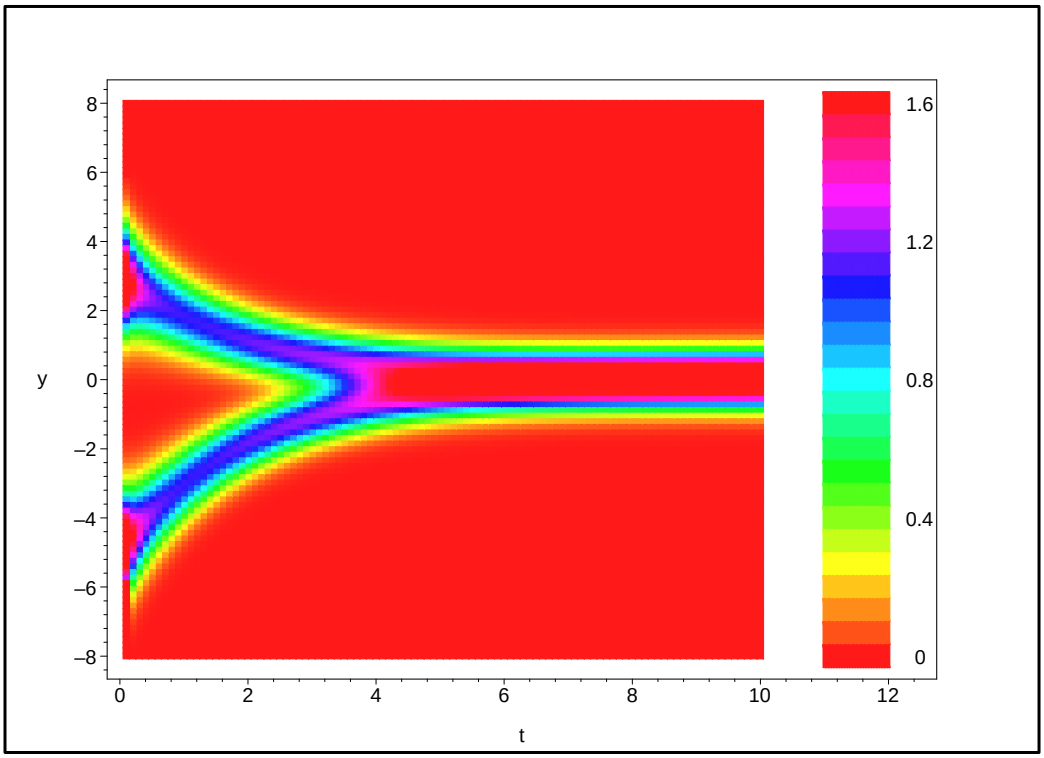


Figure 5.3: The path integration solutions $u_t = PI_t u_0$ of the Langevin equation with initial density u_0 given in (5.21) and parameters $a = b = 1/2$.

namely two Gaussian densities with unit variance and means $x = 3, x = -5$. The resulting densities $u_t = PI_t u_0$ for $t \in [0, 10]$ with parameters $a = b = 1/2$ are plotted in figure 5.3. The actual computation was performed using MAPLE. The path integration densities are easily obtained since they are the convolution integrals between Gaussian densities, but the explicit expression is omitted since it is somewhat complicated.

We have now developed the notion of path integration with exact transition kernels. At this point it is clear that if we have the exact transition kernel at hand, our problem is generally solved for any initial density u_0. However there are only a few examples where the exact kernel is known, and in general some sort of approximation procedure has to be applied.

5.3 Time-Discrete Transition Kernels

The most obvious candidate for finding approximate transition kernels is time-discretization of the underlying stochastic differential equation. The most basic discretization procedure is the Euler method which we discussed in section 3.3.

The idea is to approximate the continuous-time Markov process X_t which is the solution of our stochastic differential equation by a Markov chain $\bar{X}_{t_i}$ where $t_i = i\tau$, $i = 0, \ldots, K$. For obvious reasons we denote τ the time step. Throughout this text we will only consider the Euler scheme as our discretization procedure since it is the only scheme that, to our knowledge, has been thoroughly explored for a general Lévy-driven stochastic differential equation. For Brownian motion driven stochastic differential equations, many more schemes are presented in Kloeden & Platen (1999), and Milstein type schemes for some subclasses of Lévy-processes are described in Yan (2005)
The Euler scheme for the 1-dimensional stochastic differential equation

$$\mathrm{d}X_t = f_1(X_{t-})\mathrm{d}t + f_2(X_{t-})\mathrm{d}L_t \tag{5.22}$$

is given as

$$\bar{X}_{t_0} = x \tag{5.23}$$
$$\bar{X}_{t_i} = \bar{X}_{t_{i-1}} + f_1(\bar{X}_{t_{i-1}})\tau + f_2(\bar{X}_{t_{i-1}})[L_{t_i} - L_{t_{i-1}}]. \tag{5.24}$$

We denote the time-discrete transition kernel $\bar{k}(y, x, \tau)$. Setting $x_i = \bar{X}_{t_i}$, we obtain the characteristic function of the time discrete forward transition kernel conditioned on x_{i-1}:

Result 5.2: *The characteristic function for an increment of the Euler-discretized stochastic differential equation (5.22) is given as*

$$\mathscr{F}_{x_i}\bar{k}(x_i, x_{i-1}, \tau) = \psi_\tau(u) = \exp(iu(x_{i-1} + f_1(x_{i-1})\tau))\phi_{L_\tau}(f_2(x_{i-1})u) \tag{5.25}$$

where ϕ_{L_τ} is the characteristic function of the Lévy-increment.

Proof.

$$\psi_\tau(u) = \mathbb{E}[\exp(iu(x_{i-1} + f_1(x_{i-1})\tau + f_2(x_{i-1})\{L_{t_i} - L_{t_{i-1}}\}))]. \tag{5.26}$$

Since the Lévy-increment $\{L_{t_i} - L_{t_{i-1}}\}$ is independent of $\mathcal{F}_{t_{i-1}}$ we can write

$$\begin{aligned}\psi_\tau(u) &= \mathbb{E}[\exp(iu(x_{i-1} + f_1(x_{i-1})\tau)]\mathbb{E}[iu(f_2(x_{i-1})\{L_{t_i} - L_{t_{i-1}}\}))] \\ &= \exp(iu(x_{i-1} + f_1(x_{i-1})\tau))\phi_\tau(f_2(x_{i-1})u)\end{aligned} \tag{5.27}$$

□

Hence we know the Fourier transform of the forward transition kernel. We continue the Langevin equation example by finding the time discrete transition kernel:

Example 5.2: Recall the Langevin equation (5.16). It is easy to see that in the framework (5.22) we have that

$$f_1(x) = -ax \tag{5.28}$$
$$f_2(x) = b \tag{5.29}$$
$$L_{t_i} - L_{t_{i-1}} = B_{t_i} - B_{t_{i-1}} \tag{5.30}$$

From stochastic analysis we know that $\mathcal{L}[B_{t_i} - B_{t_{i-1}}] = N(0, \tau)$ (see e.g. Øksendal (2003)). Moreover the characteristic function of a Gaussian variable $\mathcal{L}[Y] = N(\mu, \sigma^2)$ is given as

$$\phi_Y(u) = \exp\left(i\mu u - \frac{1}{2}\sigma^2 u^2\right) \tag{5.31}$$

If we plug this into (5.27) we get:

$$\psi_\tau(u) = \exp(iu(x_{i-1} - ax_{i-1}\tau)\exp(-\frac{1}{2}\tau b^2 u^2) = \exp(iu(x_{i-1} - ax_{i-1}\tau) - \frac{1}{2}(\tau b^2)u^2). \tag{5.32}$$

Hence the time-discrete forward transition kernel is Gaussian and is given as:

$$\bar{k}(x_i, x_{i-1}, \tau) = \frac{1}{\sqrt{2\pi}\tau b^2}\exp\left(-\frac{(x_i - (x_{i-1} - ax_{i-1}\tau))^2}{2\tau b^2}\right). \tag{5.33}$$

One important fact about the time-discrete transition kernel is that the corresponding path integration operator forms a time-discrete stochastic semi-group.

Definition 5.3: *The time-discrete path integration operator* $\bar{\mathrm{PI}}_\tau$ *is defined as*

$$u_\tau(y) = \bar{\mathrm{PI}}_\tau u_0 = \int_{\mathbb{R}} \bar{k}(y, x, \tau)u_0(x)dx. \tag{5.34}$$

$\bar{\mathrm{PI}}_\tau$ obey the properties of definition 5.2 at discrete times t_i. The proofs of properties 1-4 are identical to those in the proof of result 5.1, since the Chapman-Kolmogorov equation holds for Markov chains as well.
It is natural to ask what the relation between the time discrete and the true transition kernel is:

5.4 Convergence of the Time Discrete Transition Kernel

In this section we give some preliminary results on the convergence of the time discrete transition kernel in different senses as the time step $\tau \to 0$. This is very much an unfinished theory in the Lévy case and no theorems are given. When the driving noise is Brownian motion, some theory can be found in Bally & Talay (1996) and Bally & Talay (1995).

5.4.1 Forward Convergence

The more intuitive sense of convergence is in the forward sense. That is, keeping the second argument in the kernel fixed, and taking the kernels as univariate functions of the first argument. The following result relates the weak convergence of the discretization with the forward convergence of the transition kernel.

Result 5.3: *Let $1 > \epsilon > 0$. Suppose that*

1. *$f_1, f_2 \in C^4$ and $D^i f_1, D^i f_2 \in L^\infty$ for $i = 1, 2, 3, 4$.*
2. *The exact and time discrete transition kernels $k(y, x, \tau)$, $\bar{k}(y, x, \tau) \in C_y^4$ and $D_y^i k, D_y^i \bar{k} \in L^\infty$ for $i = 0, \ldots, 4$ for all $x \in \mathbb{R},\ \tau \in (\epsilon, 1)$.*

Then there exists a constant K_ϵ, not depending on τ such that

$$\|k(y, x, \tau) - \bar{k}(y, x, \tau)\|_{y,2} \leq \sqrt{K_\epsilon \tau (1 + |x|^4)}\ \forall\ \tau \in (\epsilon, 1). \tag{5.35}$$

Proof. The proof is based on theorem 3.6. A more accessible formulation of this theorem is found in Jacod et al. (2005), which we will use in this proof. They state that for any $g \in C^4,\ D^i g \in L^\infty$ for $i = 0, \ldots, 4$ and under assumption 1 on the stochastic differential equation we have that

$$|\mathbb{E}[g(X_\tau)] - \mathbb{E}[g(\bar{X}_1)]| \leq K\tau \|g\|_{4,\infty}(1 + |x|^4) \tag{5.36}$$

where $\|g\|_{4,\infty} = \max(\{\|D^i g\|_\infty\}_{i=0}^4)$. Fix some $\epsilon > 0$ and some initial point $x_0 \in \mathbb{R}$. This implies that our initial density $u_0(x) = \delta(x - x_0)$, where $\delta(\cdot)$ denotes the Dirac delta function. Hence

$$\begin{aligned}\left|\mathbb{E}[g(X_\tau)] - \mathbb{E}[g(\bar{X}_1)]\right| &= \Big| \int_{\mathbb{R}} g(y) \int_{\mathbb{R}} k(y, x, \tau)\delta(x - x_0)\mathrm{d}x\mathrm{d}y \\ &- \int_{\mathbb{R}} g(y) \int_{\mathbb{R}} \bar{k}(y, x, \tau)\delta(x - x_0)\mathrm{d}x\mathrm{d}y \Big| = \left| \int g(y)[k(y, x_0, \tau) - \bar{k}(y, x_0, \tau)]\mathrm{d}y \right|. \end{aligned} \tag{5.37}$$

Now set $g_{x_0,\tau}(y) = k(y, x_0, \tau) - \bar{k}(y, x_0, \tau)$. This can obviously be done due to assumption 2. Moreover set

$$\begin{aligned} K_\epsilon &= K \left(\sup_{\tau \in (\epsilon, 1)} \|g_{x_0,\tau}\|_{4,\infty} \right) \\ &\leq K \left(\sup_{\tau \in (\epsilon, 1)} \|k(y, x_0, \tau)\|_{4,\infty} + \sup_{\tau \in (\epsilon, 1)} \|\bar{k}(y, x_0, \tau)\|_{4,\infty} \right) < \infty. \end{aligned} \tag{5.38}$$

Finally we compile the expressions to get

$$\int_{\mathbb{R}} [k(y, x_0, \tau) - \bar{k}(y, x_0, \tau)]^2 \mathrm{d}y = \|k(y, x_0, \tau) - \bar{k}(y, x_0, \tau)\|_{y,2}^2 \leq K_\epsilon \tau (1 + |x_0|^4) \tag{5.39}$$

which proves the desired relation since x_0 was taken arbitrarily. □

The author realizes that this result does not show convergence in the L^2 sense since K_ϵ is not controlled as $\epsilon \to 0$. This is due to the fact that derivatives of the kernels grows as $\tau \to 0$. However for fixed ϵ, the result has some relevance. We illustrate this by continuing our Langevin equation example:

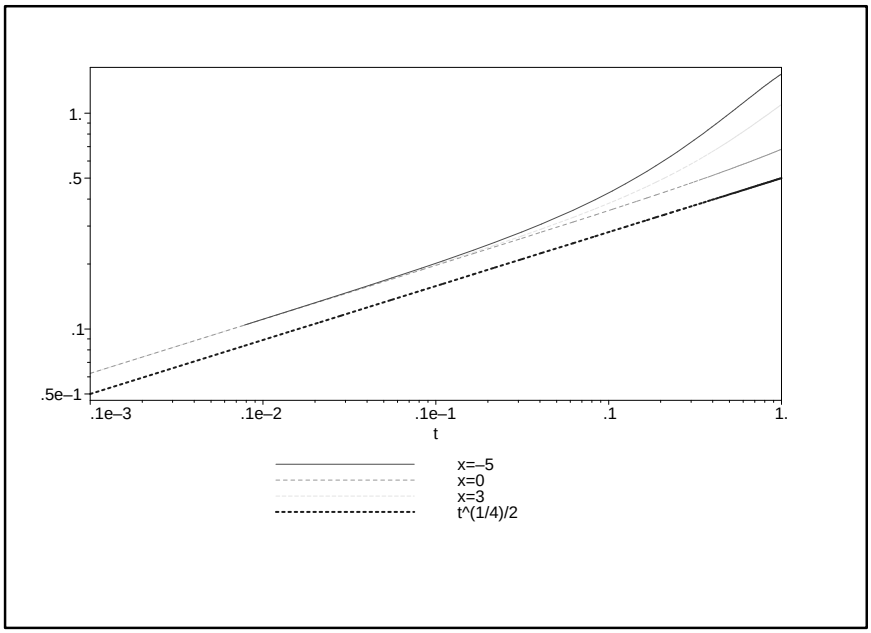


Figure 5.4: L^2-norm of the difference between the exact and time-discrete forward transition kernels for $x = -5, 0, 3$. Also included is the function $s(\tau) = \tau^{(1/4)}/2$ to illustrate what seems to be roughly the decay as τ decreases.

Example 5.3: Since we have explicit expressions for both the exact and time-discrete transition kernels, we can compute the L^2-norm of their difference. The computation is straight forward using MAPLE, but the explicit expressions are somewhat complicated and therefore omitted. Figure 5.4 shows that for $\epsilon = 0.0001$ the L^2-norm of the difference seem to decay asymptotically as $O(\tau^{(1/4)})$ for various x.

5.4.2 Backward convergence

The backward convergence is of more practical relevance since it is a key element of the convergence proof of the numerical path integration procedure which we shall present later in this chapter. In this proof we assume that

$$\lim_{\tau \to 0} \|k(y, x, \tau) - \bar{k}(y, x, \tau)\|_{x,2} = 0 \ \forall \ y \in \mathbb{R} \tag{5.40}$$

which seems reasonable under regularity constraints on f_1, f_2. No hard results are given here either, only the continuation of our Langevin equation example:

Example 5.4: In this example we explore the convergence of the backward kernels, keeping the first argument in the kernel fixed. We compute the difference norm using MAPLE. The results are presented in figure 5.5 for various values of y. Also here the decay of the norm seems to be asymptotically of $O(\tau^{(1/4)})$.

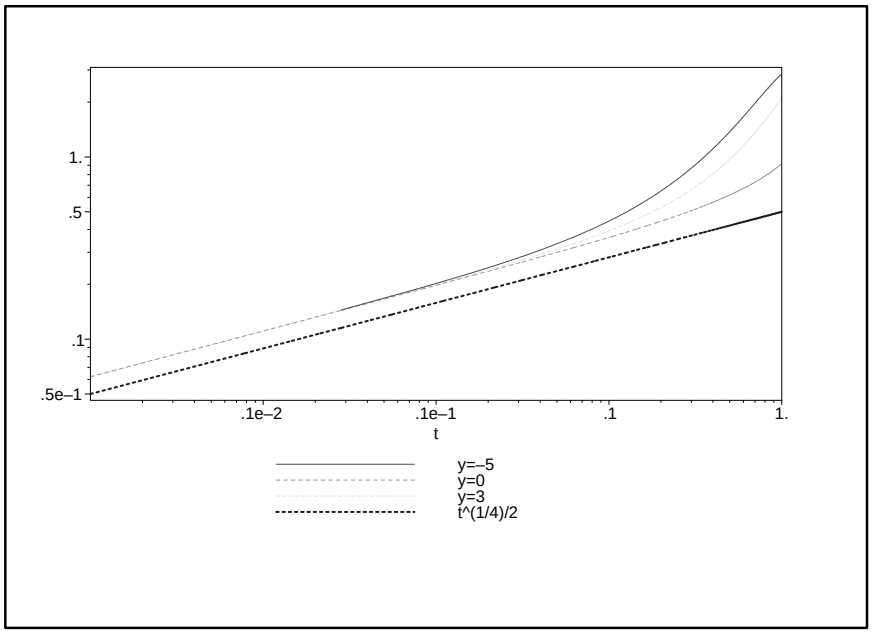


Figure 5.5: L^2-norm of the difference between the exact and time-discrete backward transition kernels for $y = -5, 0, 3$. Also included is the function $s(\tau) = \tau^{(1/4)}/2$ to illustrate what seems to be roughly the decay as τ decreases.

5.5 Numerical Path Integration

As alluded to before, one of the scopes for this chapter is to develop a numerical path integration scheme that approximates the exact path integration operator PI_t with arbitrarily accuracy. The method is well-described in the case when the driving noise is Brownian motion: see e.g. Naess (2001), Naess & Moe (2000) and references therein. From now on we denote u the numerical path integration density and u^* the exact density. We start of with some remarks concerning the application of the semi-group property of the path integration operator.

5.5.1 Evolution of Densities

The first important feature of the numerical path integration which we discuss here is the application of the semi-group property of the time discrete path integration operator. Since the time-discrete operator typically will deviate more from the exact operator as we take larger time steps, it is essential to use small time steps. However we are often interested in computing the density u_T^* at some time T which is much larger than τ. We start with a definition:

Definition 5.4: *Take some integer M and set $\tau = T/M$. We define the M-step time*

discrete PI-operator $\bar{\mathrm{PI}}_T^M$ *to be*

$$\bar{\mathrm{PI}}_T^M = \underbrace{\bar{\mathrm{PI}}_\tau(\ \bar{\mathrm{PI}}_\tau(\cdots\ \bar{\mathrm{PI}}_\tau)\cdots)}_{M\ times}. \tag{5.41}$$

Due to the Chapman-Kolmogorov equation, we can find the kernel $\bar{k}^M$ corresponding to $\bar{\mathrm{PI}}_T^M$:

$$\bar{k}^M(y,x,T) = \underbrace{\int_{\mathbb{R}}\int_{\mathbb{R}}\cdots\int_{\mathbb{R}}}_{M-1\text{ times}} \bar{k}(y,x_{M-1},\tau)\bar{k}(x_{M-1},x_{M-2},\tau) \cdots \bar{k}(x_1,x,\tau)\mathrm{d}x_1\cdots\mathrm{d}x_{M-2}\mathrm{d}x_{M-1} \tag{5.42}$$

So far so good: we have found the kernel corresponding to the M-step Euler approximation. However to do this numerically, it is common (Naess 2001) to propagate the density u trough all the steps:

$$\begin{aligned}\bar{\mathrm{PI}}_T^M u_0 &= \int_{\mathbb{R}} \bar{k}^M(y,x,T)u_0(x)\mathrm{d}x \\ &= \int_{\mathbb{R}} \bar{k}(y,x_{M-1},\tau)\Big(\int_{\mathbb{R}} \bar{k}(x_{M-1},x_{M-2},\tau)\Big(\cdots\Big(\int_{\mathbb{R}} \bar{k}(x_1,x,\tau)u_0(x)\mathrm{d}x\Big)\cdots\Big)\mathrm{d}x_{M-2}\Big)\mathrm{d}x_{M-1}\end{aligned} \tag{5.43}$$

rather then to compute the M-step transition kernel.

5.5.2 Spline Interpolation and Quadrature Integration

In order to be able to implement the numerical path integration procedure on a digital computer, we have to discretize the spatial variables at each time step. Let $\{x_i^k\}_{k=1}^Q$ be a suitable collection of equidistant nodes, where the density u_i is close to numerical zero outside this grid. We denote x_i the spatial variable corresponding to time-step i. In the computer memory we only store the values $u_{i-1}(x_{i-1}^k),\ k=1,\ldots,Q$. However, in order to integrate the old density $u_{i-1}(x_{i-1})$, we generally need to be able to evaluate it at each point in $\mathbb{R}$. To do so, we interpolate the points using some sort of spline interpolants, e.g. cubic splines (see e.g. Schultz (1973) or deBoor (2001)).
Only in a few cases is it possible to perform the actual integration of (5.43) analytically. Generally the integration has to be performed by some sort of quadrature rule on a digital computer.
Let us look a little closer on a single path integration step:

$$u_i(x_i) = \ \bar{\mathrm{PI}}_\tau u_{i-1}(x_{i-1}) = \int_{\mathbb{R}} \bar{k}(x_i,x_{i-1},\tau)u_{i-1}(x_{i-1})\mathrm{d}x_{i-1}. \tag{5.44}$$

If we set $x_i = x_i^k$, the integral becomes a scalar, which may be approximated for each k by a quadrature rule (see e.g. Burden & Faires (2000)). For a univariate continuous function

f over a compact $[a, b]$, the integral $\int_a^b f(x)\mathrm{d}x$ can be approximated with arbitrary precision by a sum:

$$\int_a^b f(x)\mathrm{d}x \approx \sum_{l=0}^{N} w_l f(x_l),\ x_l \in [a, b]. \tag{5.45}$$

The weights $\{w_l\}$ and point $\{x_l\}$ can be derived from the quadrature rule. As an example, if we take $x_l = a + l(b-a)/(N+1)$ and $w_l = (b-a)/(N+1)$, we have the so-called left endpoint rule. The integration procedures are treated in more detail in section 5.8

5.5.3 Algorithm and Further Issues of Implementation

We are now ready to give a pseudo-code for the numerical path integration procedure, assuming that the same spatial grid is used for each time step: There are many details

Algorithm 1 Numerical Path Integration

```
Initialize spatial grids
Initialize the spline matrix
LU-decompose the spline matrix
Represent the initial density u_0 as a spline function
i = 0
while i < M do
    for k ∈ {1, 2, ..., Q} do
        approximate u_{i+1}(x_{i+1}^k) by a quadrature approximation to (5.44) using the
        spline representation of u_i
    end for
    Spline represent u_{i+1}.
    i = i + 1
end while
```

to fill in in algorithm 1. The problem of choosing appropriate spatial grids: being wide enough to cover the significant part of the probability mass, but narrow enough to give reasonable resolution when the number of grid nodes Q, is reasonably small. Dr.Ing student Eirik Mo uses Monte Carlo simulation of a relatively small number paths in his codes to make inference as to where the significant part of the probability mass is situated (personal comm.). Using this information, the process of choosing spatial grids can be automated as the stochastic differential equation varies. Another procedure for choosing spatial grids would be to use thresholds on the probability mass at the tails, making the computation area wider if the density "smears" out.
Specification of the other simulation parameters (i.e. the number of grid points and time steps) is very much a tradeoff between desired accuracy and computational cost. We shall shortly explore in more detail how these parameters relates to the error in the numerical path integral approximation to u_T.

5.6 Convergence of the Numerical Path Integration Method

This section is divided into two parts. First we state some lemmas which will be useful for our convergence proof, and then state the convergence theorem.

5.6.1 Convergence Proof Lemmas

These results use the notation developed in subsection 2.2. We start of with a simple analysis result:

Lemma 5.1: *Let* $f \geq 0, f \in L^2(\mathbb{R}) \cap C^0(\mathbb{R})$ *and* $0 \leq a < b < B_f$. *Then* $\|a \wedge f\|_2 < \|b \wedge f\|_2$.

Proof. Assume first that $a > 0$. Since $f \in L^2(\mathbb{R})$ we have that $\lim_{|x|\to\infty} f(x) = 0$. Moreover due to the continuity condition both $f(x) = a$ and $f(y) = b$ have at least two solutions. Set

$$S_a = \{x; f(x) \geq a\} \tag{5.46}$$
$$S_b = \{x; f(x) \geq b\}. \tag{5.47}$$

Clearly we have the strict inclusion $S_b \subset S_a$. Define $\Delta = S_a \setminus S_b$, then

$$\begin{aligned}\|b \wedge f\|_2^2 - \|a \wedge f\|_2^2 &= b^2\mu(S_b) + \int_{\mathbb{R}\setminus S_b} f^2 - a^2\mu(S_a) - \int_{\mathbb{R}\setminus S_a} f^2 \\ &= b^2\mu(S_b) + \int_{\mathbb{R}\setminus S_a} f^2 + \int_\Delta f^2 - a^2\mu(S_b) - a^2\mu(\Delta) - \int_{\mathbb{R}\setminus S_a} f^2 \\ &= (b^2 - a^2)\mu(S_b) + \int_\Delta (f(x)^2 - a^2)\mu(\mathrm{d}x) > 0 \end{aligned} \tag{5.48}$$

The last strict inequality is obvious since $b > a$, $S_b \neq \emptyset$, and $f(x) \geq a \,\forall\, x \in \Delta$.
In the case where $a = 0$ we have that $a \wedge f = 0 \wedge f = 0$, hence $\|a \wedge f\|_2 = 0$, and the inequality $\|a \wedge f\|_2 < \|b \wedge f\|_2$ with $b > 0$ is obvious. □

From the above it is easy to see that under the same assumptions on f we have that

$$\lim_{a \searrow 0} \|a \wedge f\|_2 = 0. \tag{5.49}$$

Now a basic result from spline theory:

Lemma 5.2: *Let* $f \in C^2([L, R])$, $L, R \in \mathbb{R}$. *Then there exist a cubic spline representation* f^s *of* f *such that*

$$\|f - f^s\|_2 \leq (\Delta x)^2 K^s \sqrt{R - L} \tag{5.50}$$

where $K^s < \infty$ *and* Δx *is the maximal grid step.*

Proof. By theorem 4.5 in (Schultz 1973) we have that

$$\|f - f^s\|_2 \leq \frac{(\Delta x)^2}{(2\pi)^2}\|D^2 f\|_2 \tag{5.51}$$

where D denotes the differential-operator. Since $f \in C^2([L,R])$, $D^2 f$ is bounded by some constant $B_{D^2 f} \in \mathbb{R}^+$. Using this we approximate the right hand side of (5.50):

$$\|f - f^s\| \leq (\Delta x)^2 \frac{B_{D^2 f}}{(2\pi)^2}\sqrt{R-L}. \tag{5.52}$$

□

Now we state the first main lemma of this section:

Lemma 5.3: *Let $f, g \in \mathbb{D}$ and $f^{pt}, g^{pt} \in \mathbb{D}^{pt}$. Then there exists functions $K_1^I(L_f, R_f)$, $K_2^I(L_g, R_G)$, $K_3^I(L_f, R_f, L_g, R_g)$ with limit properties*

$$\lim_{L_f \searrow -\infty} \lim_{R_f \nearrow \infty} K_1^I = \|f\|_2 \tag{5.53}$$

$$\lim_{L_g \searrow -\infty} \lim_{R_g \nearrow \infty} K_2^I = \|g\|_2 \tag{5.54}$$

$$\lim_{L_f \searrow -\infty} \lim_{R_f \nearrow \infty} \lim_{L_g \searrow -\infty} \lim_{R_g \nearrow \infty} K_3^I = 0 \tag{5.55}$$

such that

$$\left| \int_{\mathbb{R}} f(x)g(x)\,dx - \int_{\mathbb{R}} f^{pt}(x)g^{pt}(x)\,dx \right| \leq \epsilon_g^p K_1^I + \epsilon_f^p K_2^I + \epsilon_f^p \epsilon_g^p + K_3^I = \epsilon^I \tag{5.56}$$

Proof. Define

$$p_f = f^t - f^{pt} \qquad = g^t - g^{pt} \tag{5.57}$$

$$t_f = f - f^t \qquad = g - g_t \tag{5.58}$$

It is clear that $p_f, p_g, t_f, t_g, f^t, g^t, f^{pt}, g^{pt} \in L^2(\mathbb{R})$ hence the application of L^2-innerproducts is justified:

$$\begin{aligned}\left| \int fg - \int f^{pt}g^{pt} \right| &= |\langle f,g\rangle - \langle f^{pt}, g^{pt}\rangle| = |\langle f,g\rangle - \langle f,g\rangle \\ &+ \langle f, t_g\rangle + \langle t_f, g\rangle - \langle t_f, t_g\rangle + \langle f, p_g\rangle - \langle t_f, p_g\rangle + \langle p_f, g\rangle - \langle p_f, t_g\rangle + \langle p_f, p_g\rangle| \\ &\leq \|f\|_2\epsilon_g^t + \|g\|_2\epsilon_f^t + \epsilon_f^t\epsilon_g^t + \|f\|_2\epsilon_g^p + \epsilon_f^t\epsilon_g^p + \|g\|_2\epsilon_f^p + \epsilon_f^p\epsilon_g^t + \epsilon_f^p\epsilon_g^p \end{aligned} \tag{5.59}$$

The inequality is justified by multiple applications of the Cauchy-Schwartz inequality. A final reordering yields:

$$\left| \int fg - \int f^{pt}g^{pt} \right| \leq \epsilon_g^p \underbrace{(\|f\|_2 + \epsilon_f^t)}_{K_1^I} + \epsilon_f^p \underbrace{(\|g\|_2 + \epsilon_g^t)}_{K_2^I} + \epsilon_f^p\epsilon_g^p + \underbrace{\|f\|_2\epsilon_g^t + \|g\|_2\epsilon_f^t + \epsilon_f^t\epsilon_g^t}_{K_3^I}. \tag{5.60}$$

□

Remark 5.1. The above proof is valid independently of the mutual configuration of the truncation limits.
Remark 5.2. In the case that the functions are unperturbed, the absolute value signs can be omitted, and it is clear that the truncated integral will always underestimate the true integral.

Lemma 5.4: *Let f, g, f^{pt}, g^{pt} be as in the previous lemma. Let $\mathcal{N}(f^{pt}g^{pt})$ be a quadrature approximation of $\int f^{pt}g^{pt}$ such that*

$$\left|\mathcal{N}(f^{pt}g^{pt}) - \int f^{pt}g^{pt}\right| \leq K^N(\Delta I)^\beta = \epsilon^N \tag{5.61}$$

for constants K^N, β and maximal quadrature step ΔI. Then there exist truncation limits and quadrature steps such that for each $\epsilon^N > 0$ we have that

$$\left|\mathcal{N}(f^{pt}g^{pt}) - \int fg\right| \leq \epsilon^I + \epsilon^N. \tag{5.62}$$

Proof.

$$\left|\mathcal{N}(f^{pt}g^{pt}) - \int fg\right| \leq \left|\mathcal{N}(f^{pt}g^{pt}) - \int f^{pt}g^{pt}\right| + \left|\int f^{pt}g^{pt} - \int fg\right| \leq \epsilon^N + \epsilon^I. \tag{5.63}$$

□

Now it is time to explore the path integration operator

Lemma 5.5: *Define the truncated path integration operator PI^t_τ as*

$$\mathrm{PI}^t_\tau u_i = \int_{\mathbb{R}} k^t(y, x, \tau)u_i(x)dx,\ y \in [L, R] \tag{5.64}$$

where $x \to k^t(y, x, \tau) \in \mathbb{C}^2([y - C, y + C])$, $C \in \mathbb{R}^+$ and $u_i(x) \in \mathbb{D}^{pt}([A, B])$. Then $u_{i+1}(y) = \mathrm{PI}^t_\tau u_i \in \mathbb{D}^{pt}([L, R])$.

Proof. Set

$$u_{i+1} = \mathrm{PI}^t_\tau u_i. \tag{5.65}$$

We check each property of $\mathbb{D}^{pt}([L, R])$:

1. $u_{i+1} \geq 0$: Obvious since $k, u_i \geq 0$.

2. $u_{i+1} \in L^2([L, R])$: We denote $u^*_i, u^*_{i+1} \in \mathbb{D}$ the exact time discrete densities with respect to k. Then we have that for each $y \in [L, R]$

$$|u_{i+1}(y)| \leq |u^*_{i+1}(y) + 2\epsilon^I| \tag{5.66}$$

where $\epsilon^I < \infty$ is given in (5.56). We approximate $\|u_{i+1}\|_2$ by writing

$$\|u_{i+1}\|_2 \leq \|u^*_{i+1} + 2\epsilon^I\|_2 \leq \sqrt{\|u^*_{i+1}\|_2^2 + 4(\epsilon^I + (\epsilon^I)^2)(R-L)} < \infty \tag{5.67}$$

3. $u_{i+1} \in C^0([L,R])$: First we state a small sub-lemma. Let $f(y,x)$ be continuous in the first variable for all $x \in [y-C, y+C]$, $C \in \mathbb{R}^+$ and $g(x)$ be some B_g-bounded function. Then

$$\left| \int_{y-C}^{y+C} f(y,x)g(x)\mathrm{d}x - \int_{y+h-C}^{y+h+C} f(y+h,x)g(x)\mathrm{d}x \right| \leq 2\gamma_h B_g(C-h) + 2hB_f B_g \tag{5.68}$$

where

$$\gamma_h = \sup_{x \in [y+h-C, y+C]} (|f(y,x) - f(y+h,x)|). \tag{5.69}$$

Since f is continuous in the first variable it is clear that $\gamma_h \to 0$ as $h \to 0$. Hence the integral is continuous in y when taken over compact $[y-C, y+C]$. Taken this result in account, it is clear that $u_{i+1} \in C^0([L,R])$ for each compact $[L,R]$ when we truncate $x \to k(y,x,\tau)$ symmetrically around y.

4. $u_{i+1} \in C^1([L,R])$: We simply compute Du_{i+1}:

$$\begin{aligned} Du_{i+1} = \frac{\mathrm{d}}{\mathrm{d}y} \int_{y-C}^{y+C} k(y,x,\tau)u_i(x)\mathrm{d}x = \int_{y-C}^{y+C} \frac{\partial}{\partial y} k(y,x,\tau)u_i(x)\mathrm{d}x \\ + k(y,y+C,\tau)u(y+C) - k(y,y-C,\tau)u(y-C) \end{aligned} \tag{5.70}$$

It is clear that

$$|\int_{y-C}^{y+C} \frac{\partial}{\partial y} k(y,x,\tau)u_i(x)\mathrm{d}x| \leq B_{D_1 k}\|u_i\|_1 \tag{5.71}$$

where D_1 denotes differentiation w.r.t. the first variable. Hence Du_{i+1} is finite and in view of the sub-lemma above obviously continuous since $y \to k(y,x,\tau) \in C^2$.

5. $u_{i+1} \in C^2([L,R])$: Same arguments as above when

$$\begin{aligned} D^2 u_{i+1} = \frac{\mathrm{d}^2}{\mathrm{d}y^2} \int_{y-C}^{y+C} k(y,x,\tau)u_i(x)\mathrm{d}x = \int_{y-C}^{y+C} \left(\frac{\partial^2}{\partial y^2} k\left(y,x,\tau\right) \right) u_i\left(x\right) dx \\ + 2\left(D_1\left(k\right)\left(y,y+C,\tau\right) + D_2\left(k\right)\left(y,y+C,\tau\right)\right) u_i\left(y+C\right) \\ - 2\left(D_1\left(k\right)\left(y,y-C,\tau\right) + D_2\left(k\right)\left(y,y-C,\tau\right)\right) u_i\left(y-C\right) \\ + k\left(y,y+C,\tau\right) \mathrm{D}\left(u_i\right)\left(y+C\right) - k\left(y,y-C,\tau\right) \mathrm{D}\left(u_i\right)\left(y-C\right) \end{aligned} \tag{5.72}$$

Hence $u_{i+1} \in \mathbb{D}^{pt}$.

□

Remark 5.3. From (5.72) it is clear that $D^2 u_{i-1}$ is bounded and is trivially in $L^2([L,R])$ for compact $[L,R]$.

Lemma 5.6: *Lemma 5.5 is valid when the integral is approximated by some quadrature rule when the same rule is used for all $y \in [L,R]$.*

The proof is simple when we consider a fixed integration rule: Given some integration grid $y - C \leq z_0 < z_1 < \cdots < z_N \leq y + C$, we approximate the path integration operator by

$$u_i(y) \approx \sum_{j=0}^{N} w_j k(y, z_j) u_{i-1}(z_j) \tag{5.73}$$

where $\{w_j\}_j$ is the fixed weights associated with the quadrature rule. Using this relation it trivial to show item 1-5 in the proof of lemma 5.5, and the complete proof is therefore omitted.
Now we have to address the problem of whether the path integration operator is a map $\bar{\mathrm{PI}}_t : \mathbb{D}^p \to \mathbb{D}^p$, more precisely whether the solutions stay in L^2 as the truncation limits tend to infinity. It is tempting to look for some bound on the operator norm. The classical result for a Fredholm type operator is the following:

Result 5.4: *Given the path integration operator (5.11) and assume that*

$$\|k_\tau\|_2 = \sqrt{\int_{\mathbb{R}} \int_{\mathbb{R}} |k(y,x,\tau)|^2 \, dx \, dy} < \infty. \tag{5.74}$$

Then $\|u_{i+1}\|_2 \leq \|k_\tau\|_2 \|u_i\|_2$.

Proof.

$$\begin{aligned} \|u_{i+1}\|_2^2 = \int_{\mathbb{R}} \left[\int_{\mathbb{R}} k(y,x,\tau) u_i(x) \mathrm{d}x \right]^2 \mathrm{d}y &\leq \int_{\mathbb{R}} \left[\sqrt{\int_{\mathbb{R}} |k(y,x,\tau)|^2 \mathrm{d}x} \|u_i\|_2 \right]^2 \mathrm{d}y \\ &= \|u_i\|_2^2 \int_{\mathbb{R}} \int_{\mathbb{R}} |k(y,x,\tau)|^2 \mathrm{d}x \mathrm{d}y = \|u_i\|_2^2 \|k_\tau\|_2^2 \end{aligned} \tag{5.75}$$

The inequality is justified by the Cauchy-Schwartz inequality. □

The above result shows that the operator norm of the path integration operator is bounded by $\|k_\tau\|_2$. However very little information lies in this result in our case, since $\|k_\tau\|_2 = \infty$ for most practical transition kernels. Both the exact and discretized kernels in our Langevin example have this property.
To sort this out we use the conservation of probability mass property of the stochastic semigroup.

Lemma 5.7: *The time-discrete path integration operator is a map $\bar{\mathrm{PI}}_\tau : \mathbb{D}^p(\mathbb{R}) \to \mathbb{D}^p(\mathbb{R})$.*

Proof. First we show that $\bar{\mathrm{PI}}_\tau : \mathbb{D}(\mathbb{R}) \to \mathbb{D}(\mathbb{R})$. This is clear due to the conservation of probability mass property of the stochastic semigroup, lemma 5.5 and the non-negativity of the kernel and $u_i \in \mathbb{D}$. Now assume that $u_i \in \mathbb{D}^p(\mathbb{R}) \Rightarrow u_i \geq 0$. It is clear that $\frac{1}{\|u_i\|_1} u_i \in \mathbb{D}$. Hence

$$u_{i+1} = \bar{\mathrm{PI}}_\tau u_i = \|u_i\|_1 \, \bar{\mathrm{PI}}_\tau \left(\frac{u_i}{\|u_i\|_1} \right) \tag{5.76}$$

and

$$\|u_{i+1}\|_2 = \|u_i\|_1 \left\| \bar{\mathrm{PI}}_\tau \left(\frac{u_i}{\|u_i\|_1} \right) \right\|_2 . \tag{5.77}$$

Now we know that $\left\| \bar{\mathrm{PI}}_\tau \left(\frac{u_i}{\|u_i\|_1} \right) \right\|_2 < \infty$ since $\bar{\mathrm{PI}}_\tau : \mathbb{D}(\mathbb{R}) \to \mathbb{D}(\mathbb{R})$. Hence $\bar{\mathrm{PI}}_\tau : \mathbb{D}^p(\mathbb{R}) \to \mathbb{D}^p(\mathbb{R})$. □

From the above we can conclude that the path integration procedure does not blow up (in L^2-sense) in finitely many steps.

5.6.2 Convergence Proof

Let X_t be the solution of the 1-dimensional Lévy-driven stochastic differential equation

$$\mathrm{d}X_t = f_1(X_{t-})\mathrm{d}t + f_2(X_{t-})\mathrm{d}L_t, t \geq 0, \; \mathcal{L}(X_0) = u_0. \tag{5.78}$$

We seek to approximate $u_T^* = \mathcal{L}(X_T)$ by numerical path integration for some $0 < T < \infty$. Choose $M \in \mathbb{N}$ and define $\tau = T/M$. Let u_i^s denote the ith spline represented numerical path integration density corresponding to $t_i = i\tau$ with time-discrete transition kernel $\bar{k}(y, x, \tau)$. Then we have the following theorem:

Theorem 5.1: *The L^2 numerical path integration error $\|u_T^* - u_M^s\|_2$ can be made arbitrarily small provided that:*

1. *The time discrete transition kernel has the backward convergence property*

$$\lim_{\tau \to 0} \|k(y, x, \tau) - \bar{k}(y, x, \tau)\|_{x,2} = 0 \; \forall \; y \in \mathbb{R}. \tag{5.79}$$

2. *$u_t^* \in \mathbb{D} \; \forall \; t \in [0, T]$.*
3. *u_0 is given.*
4. *All truncation limits $(L, R) \to (-\infty, \infty)$.*
5. *The spatial grid step $\Delta x \to 0$.*
6. *The quadrature grid step $\Delta I \to 0$.*

Proof. To ease notation we set $u_i^* = u_{t_i}^*$. Consider first a single numerical path integration step where the kernel is truncated according to lemma 5.5 and approximated by quadrature according to lemma 5.6. That is $x \to \bar{k}^t(y, x, \tau) = \bar{k}(y, x, \tau)\mathbf{1}_{[y-C,y+C]}$

$$u_{i+1}(x_{i+1}) = \mathcal{N}(\bar{k}^t(x_{i+1}, x_i, \tau)u_i(x_i)). \tag{5.80}$$

Define $e_i = u_i^* - u_i^s$ and $E_i = \|e_i\|_2$. By lemma 5.5 $u_{i+1} \in \mathbb{D}^{pt}[L_{i+1}, R_{i+1}]$ when $u_i \in \mathbb{D}^{pt}[L_i, R_i]$. We write

$$\|e_{i+1}\|_2 = E_{i+1} \leq \|u_{i+1}^* - u_{i+1}^{*t}\|_2 + \|u_{i+1}^{*t} - u_{i+1}\|_2 + \|u_{i+1} - u_{i+1}^s\|_2 \tag{5.81}$$

The first and last terms on the right hand side can be bounded according to lemmas 2.6, 5.2 and remark 5.3. That is, we choose L_{i+1}, R_{i+1} such that

$$\|u_{i+1}^* - u_{i+1}^{*t}\|_2 < \epsilon_{u_{i+1}^*}^t \tag{5.82}$$

and Δx such that

$$\|u_{i+1} - u_{i+1}^s\|_2 < (\Delta x)^2 K^s \sqrt{R_{i+1} - L_{i+1}} B_{D^2 u_{i+1}} = \epsilon_{i+1}^s \tag{5.83}$$

It is clear that these terms can be made arbitrarily small. The error $\bar{e}_{i+1} = u_{i+1}^{*t} - u_{i+1}$ demand some more work. It is clear that we can use lemma 5.3 to get a bound for the pointwise error, that is

$$|\bar{e}_{i+1}(y)| \leq \epsilon_{i+1}^I + \epsilon_{i+1}^N \; \forall \; y \in [L_{i+1}, R_{i+1}] \tag{5.84}$$

where

$$\epsilon_{i+1}^I = E_i(\|k\|_2 + \epsilon_k^t) + K_\tau(\|u_i\|_2 + \epsilon_{u_i}^t) + K_3^I(\epsilon_k^t, \epsilon_{u_i^*}^t) \tag{5.85}$$

$$K_\tau = \|k(y, x, \tau) - \bar{k}(y, x, \tau)\|_{x,2}. \tag{5.86}$$

However this bound will gravely overestimate the error in the tails of u^{*t}. From lemmas 5.5 and 5.7 it is clear that $u_{i+1} \in L^2(\mathbb{R})$. Since $u_{i+1}, u_{i+1}^* \geq 0$, we have that

$$|\bar{e}_{i+1}(y)| \leq F(y) = u_{i+1}(y) \vee u_{i+1}^*(y) \; \forall \; y \in \mathbb{R}. \tag{5.87}$$

This constitutes our second error bound, and we can conclude that

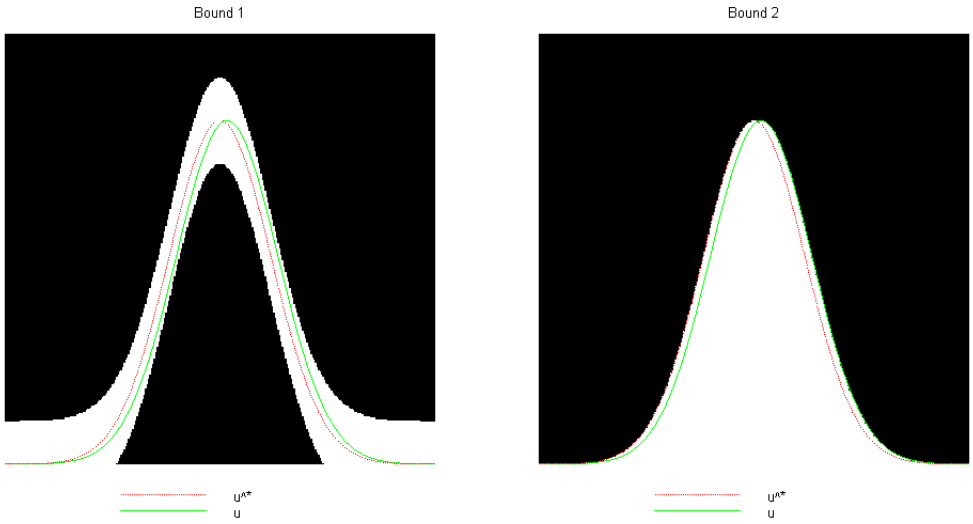
$$\|\bar{e}_{i+1}\|_2 \leq \|(\epsilon_{i+1}^I + \epsilon_{i+1}^N) \wedge F\|_2. \tag{5.88}$$

The function $\epsilon_{i+1}^I + \epsilon_{i+1}^N \wedge F$ clearly obeys the assumptions in lemma 5.1, and we have that $\|\bar{e}_{i+1}\|_2$ can be made arbitrarily small if $\epsilon_{i+1}^I + \epsilon_{i+1}^N$ can be made arbitrarily small. We compile the preliminary results to get

$$E_{i+1} \leq \epsilon_{u_{i+1}^*}^t + \epsilon_{i+1}^s + \|(\epsilon_{i+1}^I + \epsilon_{i+1}^N) \wedge F\|_2 \tag{5.89}$$

where $\epsilon_{i+1}^I \to 0$ as $K_\tau, E_i, \epsilon_k^t, \epsilon_{u_i^*}^t \to 0$. Assuming analytical initial density we have that $E_0 = 0$ and the bounding recursion (5.89) can be made arbitrarily small. □

The bounds (5.84) and (5.87) are illustrated in figures 5.6a and 5.6b. These results are assumed to generalize to n dimensions, but time constraints have prevented further work in this direction.



(a) Illustration of the first bound (b) Illustration of the second bound

Figure 5.6: (a) illustrates the pointwise bound (5.84) as a function space ball around u_i^*. u_i is guaranteed to lie inside this ball- corresponding to the unshaded part of the figure. (b) illustrates the pointwise bound (5.87) on $\bar{e}$, where $\bar{e}$ is guaranteed to lie below F. (5.87) corresponds to the unshaded part of the figure.

5.7 Numerical Path Integration for Stationary Densities

This section is devoted to stationary densities, and we work out some results relating the exact stationary density with the numerical path integration stationary density.

5.7.1 Stationary Densities

In many applications of stochastic processes, The primary interest lies in the properties of the process when it has reached a stationary state. That is when the law of the process is an eigenfunction of the path integration operator with eigenvalue 1.

Definition 5.5: *We denote the eigenfunction $\pi \in L^1$ satisfying*

$$\pi = \mathrm{PI}_t \pi \; \forall \; t \geq 0 \tag{5.90}$$

a stationary density.

We continue our Langevin example to illustrate, since it is rather easy to derive the stationary density:

Example 5.5: Recall the exact transition kernel

$$k(y,x,t) = \frac{1}{\sqrt{2\pi\frac{b^2}{2a}(1-\exp(-2at))}} \exp\left(-\frac{(y-x\exp(-at))^2}{\frac{b^2}{a}(1-\exp(-2at))}\right). \tag{5.91}$$

Letting $t \to \infty$ we obtain the asymptotic density

$$d(y) = \frac{1}{\sqrt{\pi\frac{b^2}{a}}} \exp\left(-\frac{y^2}{\frac{b^2}{a}}\right), \tag{5.92}$$

namely a Gaussian density $\mu = 0$ and $\sigma^2 = \frac{b^2}{2a}$. It is relatively easy to see that

$$d = \mathsf{PI}_t d \ \ \forall \ \ t \geq 0. \tag{5.93}$$

This teaches us an important heuristic lesson. The stationary density, if it exists, can often be found by looking at $\lim_{t\to\infty} \mathsf{PI}_t u_0$ for any initial density u_0.

From definition 5.5 it is clear that a stationary density is a fixed point of the path integration operator $\mathrm{PI}_t : L^1 \to L^1$. Since L^1 is a Banach space - it might seem reasonable to apply the Banach fixed point theorem:

Definition 5.6: *Let T be an operator $T : X \to X$ where X is a normed space. Then T is said to be a contraction if there exist some number $c < 1$ such that*

$$\|Tx - Ty\| \leq c\|x - y\| \ \forall \ x, y \in X \tag{5.94}$$

Theorem 5.2 (Banach Fixed Point Theorem)**:** *Let X be a Banach space and let $T : X \to X$ be a contraction. Then there exist one and only one fixed point.*

Proof. See e.g. theorem 9.23 in Rudin (1976). □

Corollary 5.1: *If the path integration operator is a contraction map on L^1 (or a closed subspace of L^1), it has a unique stationary density.*

Proof. Follows immediately from theorem 5.2 □

However these results are very much of theoretical interest since we generally do not have an explicit expression for the path integration operator at hand. We proceed by using the stochastic semigroup property of the path integration operator.

5.7.2 Stationary Densities of Stochastic Semigroups

We give the theorem given in Lasota & Mackey (1994), slightly rewritten:

Theorem 5.3 (Theorem 7.4.1 in Lasota & Mackey (1994))**:** *Let $\{P_t\}_{t\geq 0}$ be a stochastic semigroup of operators $P_t : L^1 \to L^1$. Assume that there is an $h \in L^1,\ h(x) \geq 0,\ \|h\|_1 > 0$ such that*

$$\lim_{t\to\infty} \|(P_t f - h)^-\|_1 = 0 \text{ for every } f \in \mathcal{D}. \tag{5.95}$$

Then there is a unique density π such that $P_t\pi = \pi\ \forall\ t \geq 0$. Furthermore

$$\lim_{t\to\infty} P_t f = \pi \text{ for each } f \in \mathcal{D}. \tag{5.96}$$

Recall example 5.5. It is clear that for any $0 < \alpha \leq 1$, $h(x) = \alpha d(x)$ does the job when $\mathcal{D}$ is the family of Dirac delta functions. This family can be extended further. It is e.g. relatively easy to see that $h(x) = \alpha d(x)$ does the job when $\mathcal{D}$ is the family of all Gaussian densities, and that $\pi = d$

5.7.3 Stationary Densities of Time-Discrete Path Integration Operators

In this subsection we discuss the relation between the stationary densities of the path integration operator and its time-discrete counterpart.

Definition 5.7: *We say that $\bar{\pi}_\tau$ is a stationary density of the time discrete path integration operator if*

$$\bar{\pi}_\tau = \bar{\mathrm{PI}}_\tau \bar{\pi}_\tau. \tag{5.97}$$

We address first the case when the path integration operator is a contraction for some fixed t, i.e. we assume the existence of $\pi = \mathrm{PI}_t\pi$. The M-fold composite time discrete path integration operator $\bar{\mathrm{PI}}_\tau^M$ is defined completely analogous to the exact operator PI_τ^M. Moreover we define the error operator

$$E_\tau f = \int e_\tau(y,x) f(x) \mathrm{d}x \tag{5.98}$$

where $e_\tau(y,x) = \bar{k}(y,x,\tau) - k(y,x,\tau)$. We use the notation E_τ^M for the M-fold composite error operator.

Result 5.5: *Let PI_t be a contraction operator on $L^1(\mathbb{R})$. Set $\tau = t/M$. Then there exists $M < \infty$ such that the M-fold time discrete operator $\bar{\mathrm{PI}}_\tau^M$ is a contraction provided that $\lim_{\tau\to 0} \|e_\tau(y,x)\|_{1,\infty} = 0$.*

Before we prove this recall that the mixed norm $\|f(x_1,x_2)\|_{1,\infty}$ is defined as

$$\|f(x_1,x_2)\|_{1,\infty} = \inf\{C \geq 0; \mu(\{x_2; \int |f(x_1,x_2)|\mathrm{d}x_1 > C\}) = 0\}. \tag{5.99}$$

This is not a very strong assumption since $k, \bar{k}$ are probability densities in the first argument.

Proof of result 5.5: Assume $f, g \in L^1(\mathbb{R})$.

$$\| \bar{\mathrm{PI}}_\tau^M f - \bar{\mathrm{PI}}_\tau^M g\|_1 = \|\mathrm{PI}_t f - \mathrm{PI}_t g + E_\tau^M f - E_\tau^M g\|_1 \tag{5.100a}$$
$$\leq \|\mathrm{PI}_t f - \mathrm{PI}_t g\|_1 + \|E_\tau^M f - E_\tau^M g\|_1 \tag{5.100b}$$

Now

$$\|E_\tau^M f - E_\tau^M g\|_1 =$$
$$\int | \int e_\tau(x_M, x_{M-1}) \int e_\tau(x_{M-1}, x_{M-2}) \cdots \int e_\tau(x_1, x_0)(f(x_0) - g(x_0))\mathrm{d}x_0 \cdots \mathrm{d}x_{M-1}|\mathrm{d}x_M$$
$$\leq \int \int |e_\tau(x_M, x_{M-1})| \int |e_\tau(x_{M-1}, x_{M-2})| \cdots \int |e_\tau(x_1, x_0)||(f(x_0) - g(x_0))|\mathrm{d}x_0 \cdots \mathrm{d}x_{M-1}\mathrm{d}x_M$$
$$\leq (\|e_\tau\|_{1,\infty})^M \|f - g\|_1 \tag{5.101}$$

and $\|\mathrm{PI}_t f - \mathrm{PI}_t g\|_1 \leq c\|f - g\|_1,\ c < 1$ by assumption. The last inequality is justified by Hölder's inequality with $p = 1,\ q = \infty$ (see e.g. Stroock (1998)). Hence

$$\| \bar{\mathrm{PI}}_\tau^M f - \bar{\mathrm{PI}}_\tau^M g\|_1 \leq [c + (\|e_\tau\|_{1,\infty})^M]\|f - g\|_1. \tag{5.102}$$

Due to the limit property of $\|e_\tau\|_{1,\infty}$ we have the desired relation since there exist M such that $(\|e_\tau\|_{1,\infty})^M < 1 - c$. □

Remark 5.4. Notice that the converse is also true, that is, if $\bar{\mathrm{PI}}_\tau^M$ can be shown to be a contraction operator, then PI_t is also a contraction operator under the assumptions in result 5.5. To see this; redo the proof with the error kernel $\hat{e}_\tau(y, x) = -e_\tau(y, x) = k(y, x, \tau) - \bar{k}(y, x, \tau)$. We have the obvious equality $\|e_\tau\|_{1,\infty} = \|\hat{e}_\tau\|_{1,\infty}$.

Now we turn to relation between the stationary densities of PI_τ and $\bar{\mathrm{PI}}_\tau$. As mentioned before, in many applications such as mechanics and numerical solution of integro-partial differential equations we are primarily interested in the stationary solutions. We give an abstract bound for the two densities:

Result 5.6: *Assume that* $\lim_{M\to\infty} \bar{\mathrm{PI}}_\tau^M \pi = \bar{\pi}_\tau$ *for every* $\tau > 0$. *Then* $\|\pi - \bar{\pi}_\tau\|_1$ *can be made arbitrarily small if* $\lim_{\tau\to 0} \|e_\tau(y, x)\|_{1,\infty} = 0$.

Proof. Fix some τ and choose M such that $\| \bar{\mathrm{PI}}_\tau^M \pi - \bar{\pi}_\tau\| < \epsilon_\tau$. Then

$$\|\pi - \bar{\pi}_\tau\|_1 \leq \|\pi - \bar{\mathrm{PI}}_\tau^M \pi\|_1 + \| \bar{\mathrm{PI}}_\tau^M \pi - \bar{\pi}_\tau\| \tag{5.103a}$$
$$\leq \|\pi - \pi - E_\tau^M \pi\|_1 + \epsilon_\tau \tag{5.103b}$$
$$\leq \|\pi\|_1 (\|e_\tau\|_{1,\infty})^M + \epsilon_\tau \tag{5.103c}$$
$$= (\|e_\tau\|_{1,\infty})^M + \epsilon_\tau \tag{5.103d}$$

The details are analogous to those in the proof of result 5.5 □

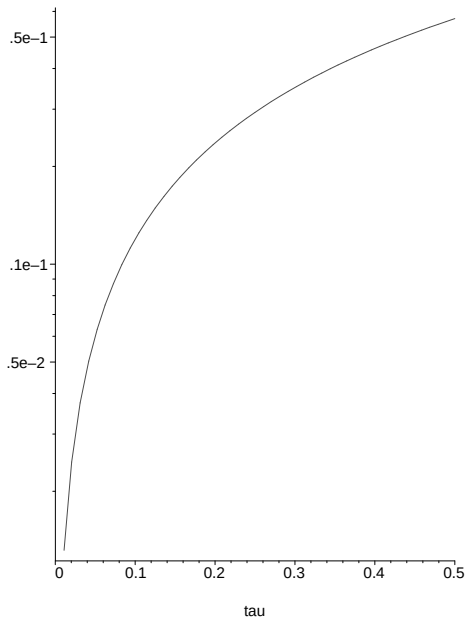


Figure 5.7: The L^1 difference between the exact and time discrete stationary densities of the Langevin equation with $a = 1/2$.

No explicit error bounds can be found using this technique, but the rate of convergence can generally be estimated using some class of initial densities believed to contain the stationary density.
To illustrate further the exact/time discrete stationary densities, we look once more at the Langevin equation:

Example 5.6: It can be shown (Kolnes 2004) that the time discrete path integration method with Euler discretization applied to the Langevin equation has a stationary density given by:

$$\bar{\pi}_\tau(y) = \frac{1}{\sqrt{\pi \frac{b^2}{a+a^2\tau/2}}} \exp\left(-\frac{y^2}{\frac{b^2}{a+a^2\tau/2}}\right). \tag{5.104}$$

We can compute the difference in stationary densities

$$\|\pi - \bar{\pi}_\tau\|_1 = 2\mathrm{erf}\left(\frac{\sqrt{2+a\tau}\sqrt{-\ln(2)+\ln(2+a\tau)}}{\sqrt{2\tau a}}\right) - 2\mathrm{erf}\left(\frac{\sqrt{-\ln(2)+\ln(2+a\tau)}}{\sqrt{\tau a}}\right), \tag{5.105}$$

where erf denotes the error function defined as

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\mathrm{d}t. \tag{5.106}$$

$\|\pi - \bar{\pi}_\tau\|_1$ is easily seen to converge and is plotted with $a = 1/2$ in figure 5.7.

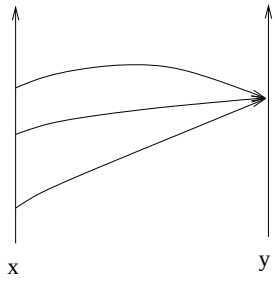


Figure 5.8: Illustration of the direct integration method. We integrate over the x axis, keeping y fixed in some grid node. Using this method demands that the transition probability is given by $\bar{k}$ is given for each x.

5.8 Time Stepping Procedures

This section describes three methods for actually evaluating the time discrete path integration integral

$$u_{i+1}(y) = \int \bar{k}(y, x, \tau) u_i(x) \mathrm{d}x. \tag{5.107}$$

The choice of method depends very much on the information about the time-discrete transition kernel given for any concrete problem. Throughout this section we omit the bars on the time-discrete path-wise variables to ease the notation. Moreover we look at stochastic differential equations on the form

$$\mathrm{d}X_t = f_1(X_t)\mathrm{d}t + f_2(X_t)\mathrm{d}L_t \tag{5.108}$$

and use the standard Euler scheme and time discretization described before.

5.8.1 Direct Integration

The first time stepping method is the direct integration method. Having the complete time discrete transition kernel $\bar{k}$ at hand we approximate the integral (5.107) with a quadrature rule and truncation (see figure 5.8). This is the most obvious way of doing the time-stepping, but for it to be applicable, we need to know the transition kernel which is not always the case. Another downside is that using this method we embed the τ-dependent stability (we will look at numerical stability shortly) of the forward Euler-scheme, forcing us to take short time-steps when the stochastic differential equation in question has a stiff nature.

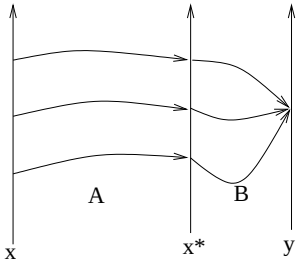


Figure 5.9: Illustration of the split step forward integration method. We integrate over the x^* axis, keeping y fixed in some grid node. Being inefficient implementation-wise, this method is not used in practice.

5.8.2 Split Step Forward Integration

The next method splits the time-discrete scheme into two stages, A and B. The first stage is the deterministic flow A given as

$$X^*_{i+1} = X_i + f_1(X_i)\tau. \tag{5.109}$$

The second stage is the stochastic increment given as

$$X_{i+1} = X^*_{i+1} + f_2(X_i)\Delta L_i \tag{5.110}$$

where $\Delta L_i = L_{t_{i+1}} - L_{t_i}$. Using this method we integrate over the x^*-space, needing only to know the density of the increment ΔL times a constant $f_2(X_i)$. This method is illustrated in figure 5.9. If we have only given the density of ΔL on a discrete grid, which is the case when we use iFFT-computed densities, this method is works poorly since one generally have to solve an implicit equation for each evaluation of the old density u_i. Due to this, this method not used in practice in this text.

5.8.3 Split Step Backward Integration

This method uses the so-called split step backward Euler scheme (see (Higham, Mao & Stuart 2002)). It can be thought of as the Euler scheme applied to a slightly altered stochastic differential equation, which converges with the original stochastic differential equation as $\tau \to 0$ (Higham et al. 2002).
This is also a two stage scheme, consisting of an implicit step for the deterministic part (A) and the stochastic part (B) treated consecutively. The determining equation for step A is given as

$$X^*_{i+1} = X_i + f_1(X^*_{i+1})\tau \Leftrightarrow X_i = X^*_{i+1} - f_1(X^*_{i+1})\tau \tag{5.111}$$

and step B given as

$$X_{i+1} = X^*_{i+1} + f_2(X^*_{i+1})\Delta L_i. \tag{5.112}$$

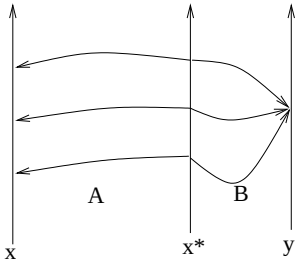


Figure 5.10: Illustration of the split step backward integration method. We integrate over the x^* axis, keeping y fixed in some grid node.

In the implementation we choose a grid, typically centered around y, in the x^*-space. Using the right hand side of (5.111), we have an explicit formula for the corresponding points in x-space, which we use to determine where to evaluate the old density. Finally we add the noise in step B, using only the density of the increment multiplied by a constant. In the case when we have additive noise, i.e. $f_2(x) = c$, this method is very efficient, since we need only compute the density of $c\Delta L$ once. This method is illustrated in figure 5.10.
The split step backward integral is easily seen to be

$$u_{i+1}(y) = \int k^*(x^*)u_i(x^* - f_1(x^*))[1 - Df_1(x^*)]\mathrm{d}x^* \tag{5.113}$$

where the measure correction factor $[1 - Df_1(x^*)]$ is needed since the spatial measure on the x-space is deformed due to (5.111). It is relatively easy to see that $k^*(x^*) = d(-x^*)$ where $d = \mathcal{L}[f_2(x^*)\Delta L]$. In the case of additive noise, we have that $d = \mathcal{L}[c\Delta L]$, i.e. a fixed probability density. This method is especially suitable for using the iFFT computed forward transition kernels when we have additive noise. The approximate kernels are computed over the x^*-space. In the code written for additive noise stochastic differential equations, we invert $\phi_{L_\tau}(-cu)$ to obtain the x^*-spatial reflection. Moreover we discard half the values computed (a quarter of the points in each tail), to reduce the relative error introduced by the iFFT computations. In the examples in the next chapter, this method is used on the additive noise equations.

5.8.4 A Brief Note on Stability

As in numerical analysis of ordinary differential equations, stability is a key attribute to any numerical scheme. The path integration schemes inherit the stability of the underlying time-discreteization scheme. Some theory on stability can be found in Kloeden & Platen (1999) and Milstein, Platen & Schurz (1998) for Brownian motion driven equations. To our knowledge very little is written about stability for general Levy-driven stochastic differential equations. The inclusion of this subsection is therefore only a heuristic comparison of stability properties of the direct integration and split step

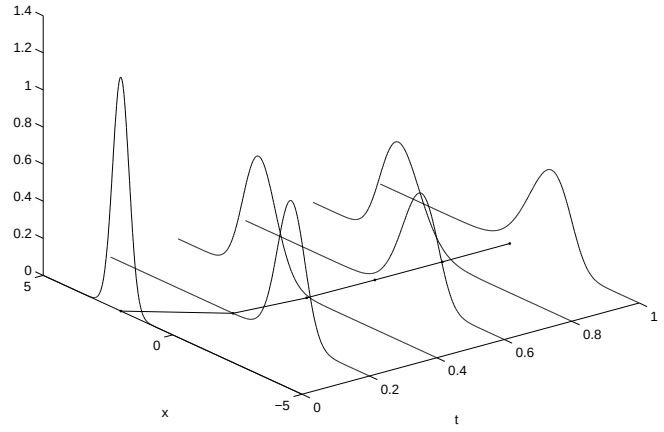


Figure 5.11: Five steps with direct integration operator with $\tau = 0.2$. The theoretic mean is also included. Notice the unstable properties that generalize those of an unstable numerical solution of an ordinary differential equation

backward integration methods. The direct integration method is based on the explicit Euler-scheme, whereas the split step backward integration method is based on the semi-implicit split step Euler scheme. From numerics on ordinary differential equations, we know that implicit methods generally show better stability properties than explicit.
There is no standard formalism on the notion of stiff stochastic differential equations (Milstein et al. 1998). In this text, we use a generalization of the notion of stiff ordinary differential equations. That is, we say a stochastic differential equation is stiff if the ordinary differential equation that arises when we set $f_2 = 0$ is stiff. We illustrate this with an example, comparing the direct integration method with split step backward integration method. Again the Langevin method is under consideration.

Example 5.7: Consider the Langevin equation

$$\mathrm{d}X_t = -10X_t\mathrm{d}t + b\mathrm{d}B_t. \tag{5.114}$$

If we let $b \to 0$ we obtain a stiff ordinary differential equation, hence we say that (5.114) is a stiff stochastic differential equation. From now on we set $b = 1$. Kloeden & Platen (1999) shows that the Euler scheme is stable is $\tau < 0.2$. If we apply the direct integration scheme on (5.114) with $\tau = 0.2$ we get the results given in figure 5.11. We see that the path integration solution contradicts the behavior we have gained intuitively through this chapter. If we use the split step backward integration method, we get the results presented in figure 5.12. We see that the computation is much more stable, supporting the hypothesis that the split step backward integration method has better stability properties.

As this example implies, the split step backward integration method is thought to have

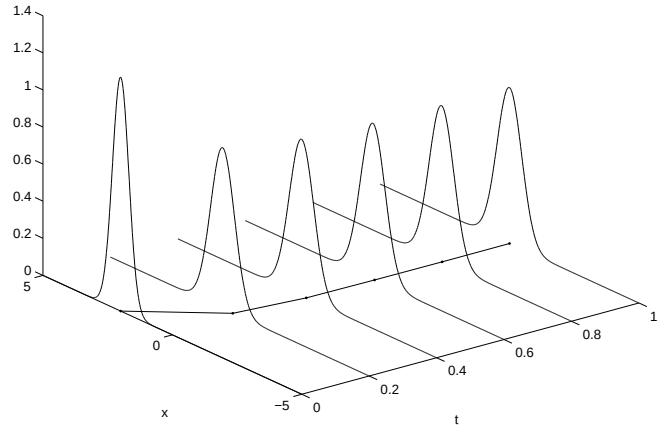


Figure 5.12: Five steps with a split step backward integration operator with $\tau = 0.2$. The theoretic mean is also included. Notice the stability compared with the direct integration operator.

better stability properties. A striking difference between ordinary and stochastic numerics is however that the computational loads of two methods are approximately the same. This is a clear argument for using the split step backward integration method in most cases.

5.9 Concluding Remarks

Throughout this chapter we have worked trough stochastic dynamics in a path integration perspective. Moreover, we have seen that the time discrete path integration method can be applied to approximate the path integration operator with arbitrary accuracy when the transition kernels converge in various manners. In the next chapter we apply these results on some real problems.

Chapter 6

Numerical Examples

In this chapter we look at numerical examples of time discrete path integration applied on various problems. Some of the `Fortran 90` codes used in this chapter can be found in appendix B. All computations were performed on a `Sun Fire V490 (sylow.math.ntnu.no)` server with 4 CPUs using the `Sun Studio 10 Fortran 95` compiler. The spline library used in the computations is (Burkardt 1999). We use B-splines (deBoor 2001) in all the examples. The iFFT routine is (*Sun Studio 11: Sun Performance Library Reference Manual - ZFFTD* 2005). The chapter is laid out as follows: First we look at the Langevin equation with various driving noises. Then Black-Scholes type equations are treated. Finally we look at a nonlinear equation.

6.1 Path Integration for the Langevin Equation

In this section we look at the Langevin equation for various driving processes. In each case we derive the Kolmogorov forward equation, which the numerical path integration method solves approximately. Stringent derivations of the Kolmogorov equations can be found in Hanson (2006) for the jump diffusion processes used below. In all the Langevin equation codes we use the split step backward operator.

6.1.1 The Langevin Equation driven by Brownian Motion

Throughout the previous chapter we worked out many of the properties of the law of the solution of the Langevin equation driven by Brownian motion. We can use these results to measure the performance of the numerical path integration procedure also described in the previous chapter. Even though the exact time-discrete transition kernel is at hand we use the iFFT computed density so we can look at how the number of FFT-points affects the accuracy.
From example 2.1 we know that the forward Kolmogorov operator is given as

$$A^*k(y) = aD(yk(y)) + \frac{b^2}{2}D^2k(y) \tag{6.1}$$

and in the examples in the previous chapter we developed the basic solution, namely the exact transition kernel.
First we look at the stationary density. Since the stationary density is known, we

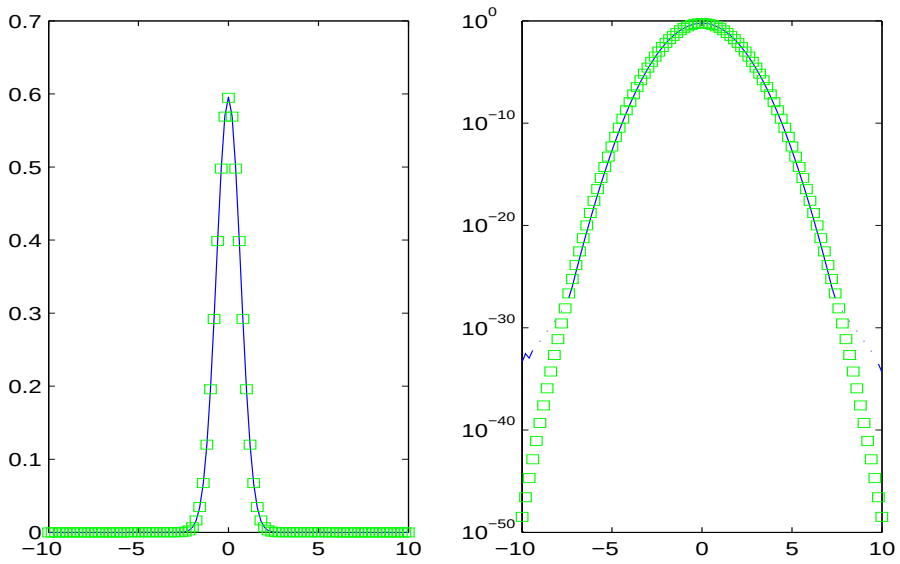


Figure 6.1: Stationary densities of the Langevin equation with $a = 0.1$, $b = 0.3$. (-) is the numerical path integration density and (□) is the exact stationary density (5.92). The numerical path integration density is computed with 101 spatial grid points, $\tau = 0.1$ and the integrals are computed with 2^4 points with 3.5 standard deviations at each side. The plots show good accordance down to approximately 10^{-25}.

start the numerical path integration with this initial density, and look for time-discrete stationary densities. The program is run until it finds an exact fixed point. When started in the stationary density, this occurs at roughly $t \approx 30$. Figure 6.1 illustrates the time-discrete versus the exact stationary density. We measure the error as

$$err = \frac{1}{Q}\sqrt{\sum_{i=1}^{Q}(u^*(x_i) - u(x_i))^2}, \tag{6.2}$$

where u^* is the exact stationary density, u the time-discrete stationary density and $\{x_i\}_{i=1}^{Q}$ the computation spatial grid.
Fixing the time step $\tau = 0.1$, we look at how the truncation of the transition probability density and the number of FFT-points affect the accuracy. The results are plotted in figure 6.2. We see that when we are working with approximate transition probabilities it is not always optimal to let the truncation limits be too wide when the relative error in the tails grows.
Next we study the error as the time step τ varies (see figure 6.3). Also here we see that effects in the numerical implementation will at some point stop us from gaining more accuracy by reducing τ when the spatial grid is kept fixed. Also included in figure 6.3 is a run using the exact transition kernel. We see that very little is gained comparing

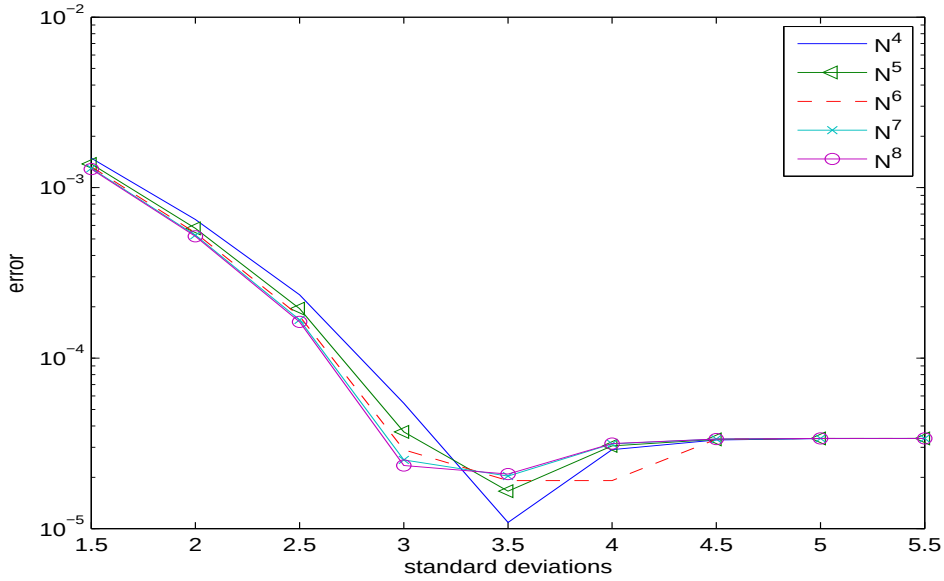


Figure 6.2: Error in the numerical path integration stationary densities for $\tau = 0.1$, $a = 0.1$, $b = 0.3$. Each line corresponds to the number of points used in the quadrature for approximating the integrals. We see that in this case, truncating at approximately 3.5 standard deviations seems close to optimal. This is explained by a balance between including the significant part of the transition probability and not including the tails where the relative error is large. Notice also that little is gained by using many points in the quadrature, admitting very fast implementations.

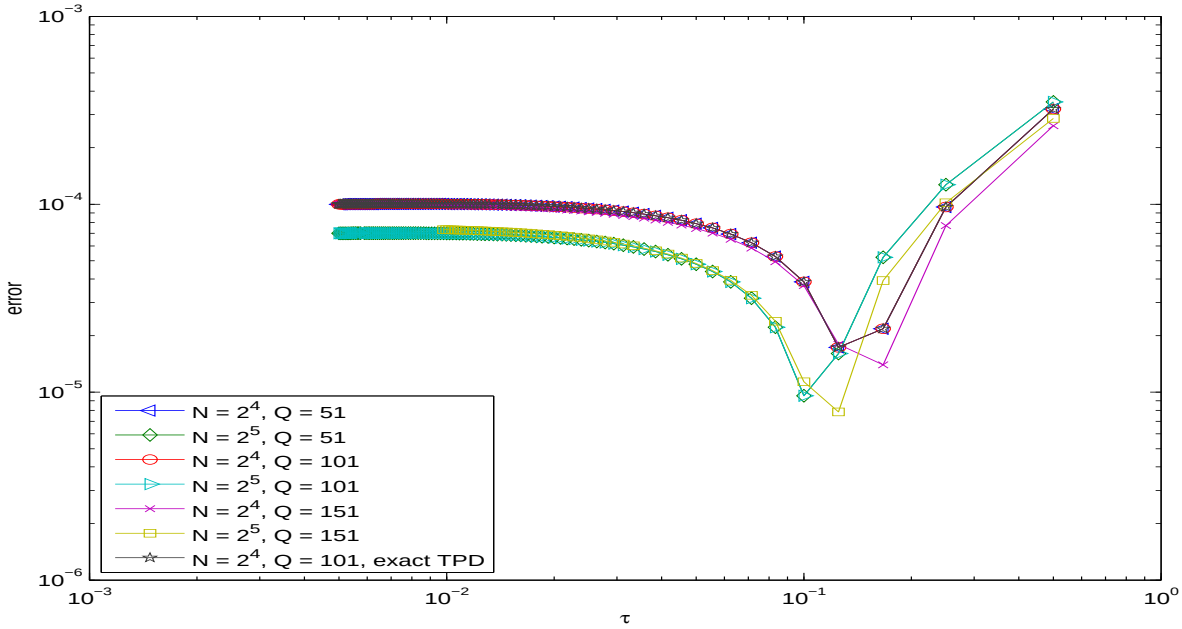


Figure 6.3: The stationary density error for various parameter settings with $a = 0.1,\ b = 0.3$. Here N denotes the number of quadrature points and Q is the number of spatial grid points. We see that the optimal time step in this case is ~ 0.1.

with the iFFT computed densities. Hence all of the increasing error effect is explained by scaling problems. I.e. problems arising when the integration interval becomes small compared to the spatial grid spacing. This amplifies the error introduced by the spline interpolation and gives poorer resolution. We see that finer grids generally gives smaller errors, supporting this observation. Hence to gain more accuracy we have to balance the time and spatial steps. The CPU-times for various parameter settings are of some interest. As always such measurements should be viewed upon as estimates since other computational and administrative processes run on the same server. The CPU-times of the runs described above are presented in figure 6.4.

Finally we include two illustrations, figures 6.5a and 6.5b, to show quantitatively how the behavior of the solution process is. We use the initial density (5.21).

6.1.2 The Langevin Equation Driven by a Compound Poisson Processes with Gaussian Jumps

Consider the 1-dimensional stochastic differential equation

$$\mathrm{d}Y_t = -aY_t\mathrm{d}t + b\mathrm{d}B_t + b\mathrm{d}X_t,\ Y_0 = y,\ a, b > 0 \tag{6.3}$$

where B_t is standard Brownian motion and X_t a compound Poisson process. Existence and uniqueness results for such equations can be found in e.g. Øksendal & Sulem (2005)

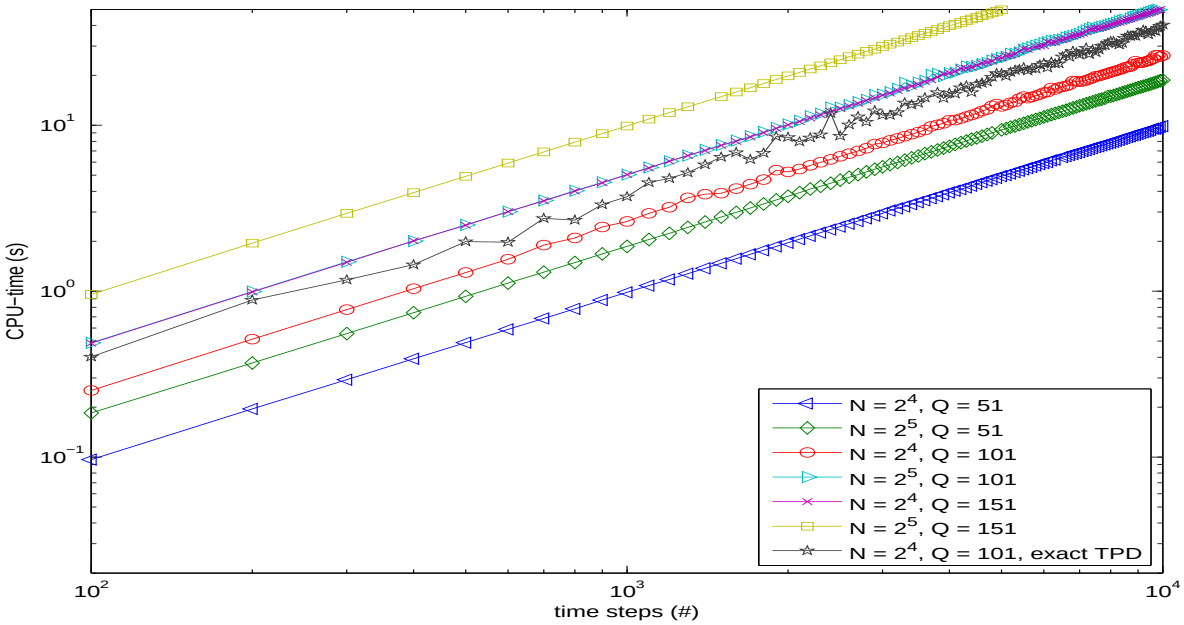
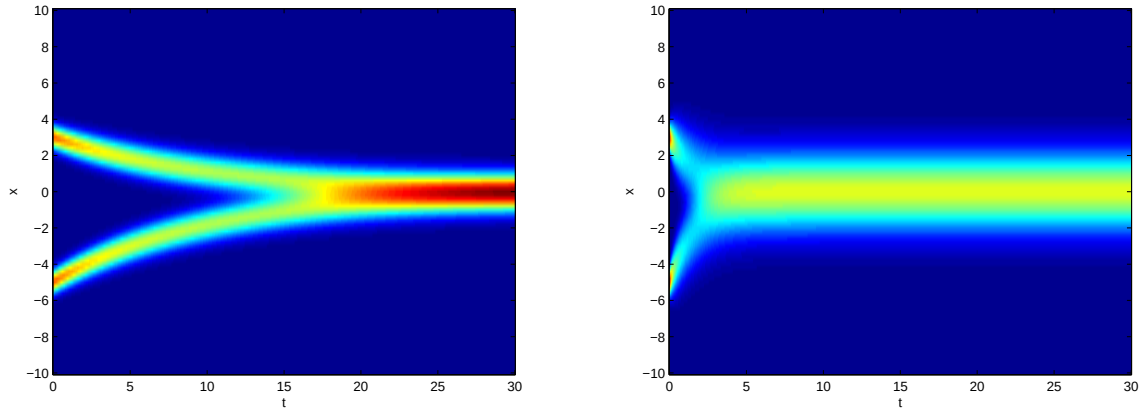


Figure 6.4: Approximate CPU-times for the Langevin equation with $a = 0.1,\ b = 0.3$. Here N denotes the number of quadrature points and Q is the number of spatial grid points. The computions were performed ut to $t = 50$.



(a) $a = 0.1,\ b = 0.3$ (b) $a = 0.5,\ b = 1.5$

Figure 6.5: Illustration of the time-evolution of the law of the solution to the Langevin equation for various parameter settings.

or Bass (2004). In this case we take the jumps to be Gaussian with zero mean. I.e $X_t = \sum_{i=1}^{N_t} Z_i$ where $\mathcal{L}[Z_i] = N(0, \sigma_Z^2)$. Let λ denote the rate of the Poisson process N_t. In our computation we collect the Levy process $U_t = B_t + X_t$. We have already computed the density and characteristic function for these increments in subsection 4.2.
It is relatively easy to see, given equation 3.24, that the infinitesimal generator of the solution process is given as

$$\begin{aligned} Ag(y) &= ayDg(y) + \frac{1}{2}b^2D^2g(y) + \int_{\mathbb{R}} g(y+bz) - g(y) - z(Dg(y))\mathbf{1}_{[-1,1]}(z)\nu(\mathrm{d}z) \\ &= ayDg(y) + \frac{1}{2}b^2D^2g(y) - \lambda g(y) + \frac{\lambda}{\sqrt{2\pi\sigma_Z^2}}\int_{\mathbb{R}} g(y+bz)\exp\left(-\frac{z^2}{2\sigma_Z^2}\right)\mathrm{d}z. \end{aligned} \tag{6.4}$$

Now it is natural to ask how the adjoint operator looks like. That is, we wish to find A^* such that

$$\int (Ag)h = \int g(A^*h) \;\; \forall\, g, h \in C_c^2. \tag{6.5}$$

It is clear that A can be split into three parts:

$$\mathcal{D}g(y) = -ayDg(y) + \frac{1}{2}b^2D^2g(y) \tag{6.6}$$

$$\mathcal{I}g(y) = \frac{\lambda}{\sqrt{2\pi\sigma_Z^2}}\int_{\mathbb{R}} g(y+bz)\exp\left(-\frac{z^2}{2\sigma_Z^2}\right)\mathrm{d}z \tag{6.7}$$

$$\mathcal{C}g(y) = -\lambda g(y) \tag{6.8}$$

and due to the linearity of the integral it is clear that $(\mathcal{D}+\mathcal{I}+\mathcal{C})^* = \mathcal{D}^* + \mathcal{I}^* + \mathcal{C}^*$. Recall from example 2.1 that the differential operator $\mathcal{D}$ has the adjoint operator

$$\mathcal{D}^*g(y) = -D(ay\, g(y)) + \frac{1}{2}b^2D^2g(y). \tag{6.9}$$

The adjoint operator for the integral operator $\mathcal{I}$ is also easily obtained using the definition. Let $g, h \in C_c^2$ and set

$$n(z) = \frac{\lambda}{\sqrt{2\pi\sigma_Z^2}}\exp\left(-\frac{z^2}{2\sigma_Z^2}\right). \tag{6.10}$$

Setting	a	b	λ	σ_Z^2	$Y_0 \sim$	τ
1	0.1	0.3	3.5	0.9^2	$N(-3, 0.1)$	1/20
2	0.2	0.8	10.5	0.3^2	$N(3, 4)$	1/20

Table 6.1: Parameter settings for the jump diffusion (6.3).

Then

$$\int (\mathcal{I}g(y))h(y)\mathrm{d}y = \int \left[\int g(y+bz)n(z)\mathrm{d}z\right] h(y)\mathrm{d}y \tag{6.11a}$$

$$= \int\int h(y)g(y+bz)n(z)\mathrm{d}y\mathrm{d}z \tag{6.11b}$$

$$= \int\int h(y-bz)g(y)n(z)\mathrm{d}y\mathrm{d}z \tag{6.11c}$$

$$= \int g(y)\left[\int h(y-bz)n(z)\mathrm{d}z\right]\mathrm{d}y \tag{6.11d}$$

$$= \int g(y)(\mathcal{I}^* h(y))\mathrm{d}y. \tag{6.11e}$$

The constant operator $\mathcal{C}$ is trivially selfadjoint, i.e $\mathcal{C} = \mathcal{C}^*$. We compile the results to find

$$A^*k(y) = -ayDk(y) + \frac{1}{2}b^2D^2k(y) - (\lambda + a)k(y) + \int k(y-bz)n(z)\mathrm{d}z \tag{6.12}$$

and our Kolmogorov forward equation becomes

$$\frac{\partial}{\partial t}k(y,x,t) = A^*_{(y)}k(y,x,t),\ y \in \mathbb{R},\ t \in (0,\infty) \tag{6.13}$$

with initial condition $\lim_{t\searrow 0} = \delta(y-x)$. The subscript (y) indicates that A^* operates on the forward kernel.

The path integration code is easy to adapt to this process. All that is needed is to change the characteristic function of the transition density. Solving the forward Kolmogorov equation analytically seems hard, but the characteristic function of the forward transition kernel is relatively easy to obtain. This is done in appendix A. Hence we use the iFFT computed density from (A.12) as reference for the numerical path integration solutions. Throughout this example we use $Q = 101$ grid points and $N = 2^7$ quadrature points. Since the forward transition kernel has semi-heavy tails, the computations are performed with $l = 20$. We do two runs with the parameters given in table 6.1. Figures 6.6, 6.7 show the path integration density along with the iFFT computed density at time $T = 10$. We see that the path integration procedure yields good results at least down to approximately 10^{-8}. For density levels below this, the picture is somewhat blurry since the relative error in the iFFT computed density grows. Finally we include plots (figures 6.8a, 6.8b) of the time evolution of $\mathcal{L}[Y_t]$, which coincides approximately with the solution of the forward Kolmogorov equation (6.13).

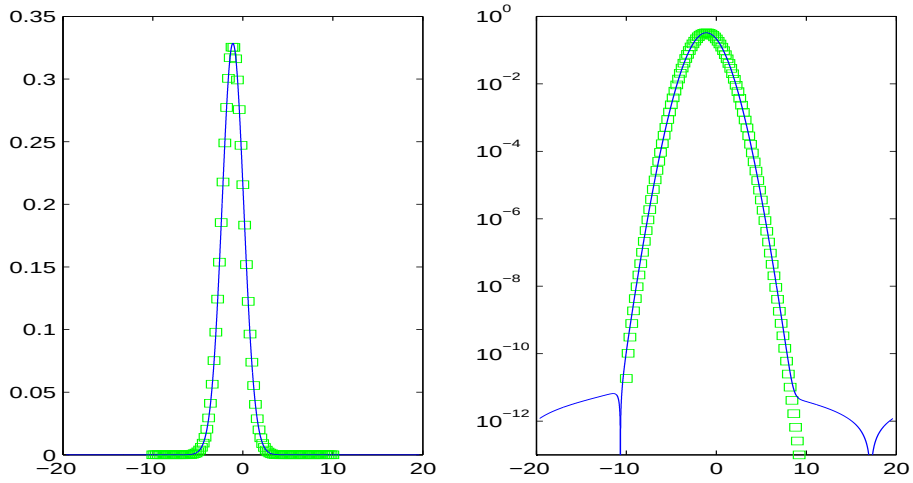


Figure 6.6: Path integration solution (□) and iFFT computed density (-) with parameter setting 1 at $t = 10$. The iFFT computation was done with $N = 2^{14}$ points.

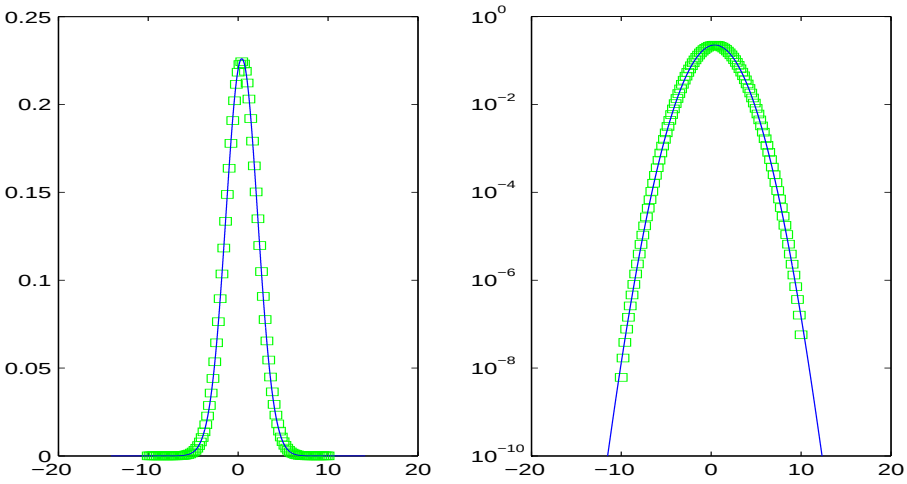
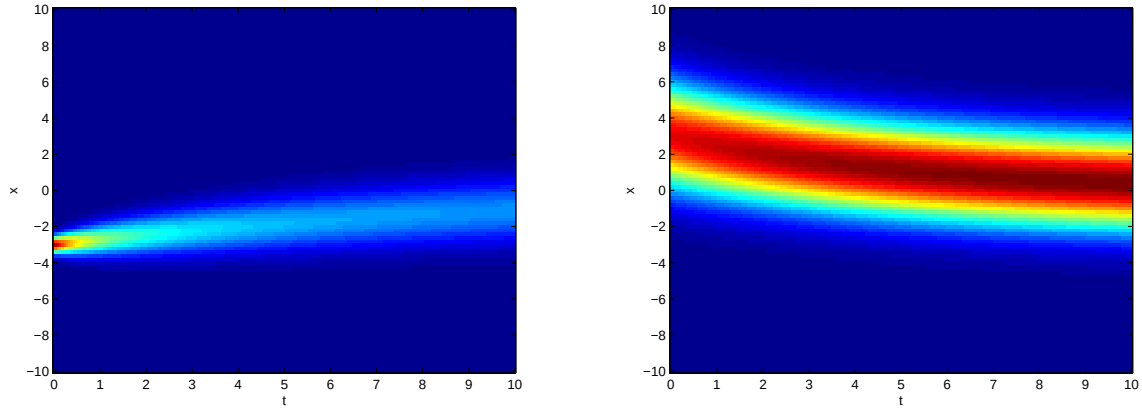


Figure 6.7: Path integration solution (□) and iFFT computed density (-) with parameter setting 2 at $t = 10$. The iFFT computation was done with $N = 2^{14}$ points.



(a) Parameter Setting 1

(b) Parameter setting 2

Figure 6.8: Illustration of the time-evolution of $\mathcal{L}[Y_t]$.

6.1.3 The Langevin Equation Driven by a Compound Poisson Processes with Exponentially Distributed Jumps

This is the last of our Langevin equation examples. Let our stochastic differential equation be given as

$$\mathrm{d}Y_t = -aY_t\mathrm{d}t + b\mathrm{d}B_t + b\mathrm{d}R_t,\ Y_0 = y,\ a, b > 0. \tag{6.14}$$

It deviates from the examples above since the compound Poisson process is in this case not a martingale. Let R_t be the usual compound Poisson process, but let $\mathcal{L}[Z_i] = EXP(1/\beta)$. We know that in this case the jumps take nonnegative values only, making the transition kernel asymmetric. We derived the characteristic function of the transition kernel in subsection 4.3.
Using the same arguments as in the previous case we see that the Kolmogorov Forward operator becomes

$$A^*k(y) = ayDk(y) + \frac{1}{2}b^2D^2k(y) - (\lambda + a)k(y) + \int k(y - bz)\mathcal{E}(z)\mathrm{d}z \tag{6.15}$$

where

$$\mathcal{E}(z) = \mathbf{1}_{[0,\infty]}(z)\frac{1}{\beta}\exp\left(-\frac{z}{\beta}\right). \tag{6.16}$$

As we saw in subsection 4.3, this transition kernel has the asymmetric tails property. Therefore the truncation of the kernel is done asymmetrically, discarding all points where $k \leq$ 1e-14, which gives better numerical results. In this case we use 2^{10} points in the iFFT, and then discarding approximately half of these as false values. Due to the heavy right hand side tail, we use $l = 50$ standard deviations in the iFFT computation of the transition kernel. We run the path integration procedure with parameter settings given

Setting	a	b	λ	β	$Y_0 \sim$	τ
1	1.1	0.9	1.5	1.1	$N(3, 1.1)$	1/20
2	0.5	0.9	1.5	1.1	$UNIF([-6, 6])$	1/20

Table 6.2: Parameter settings for the jump diffusion (6.14).

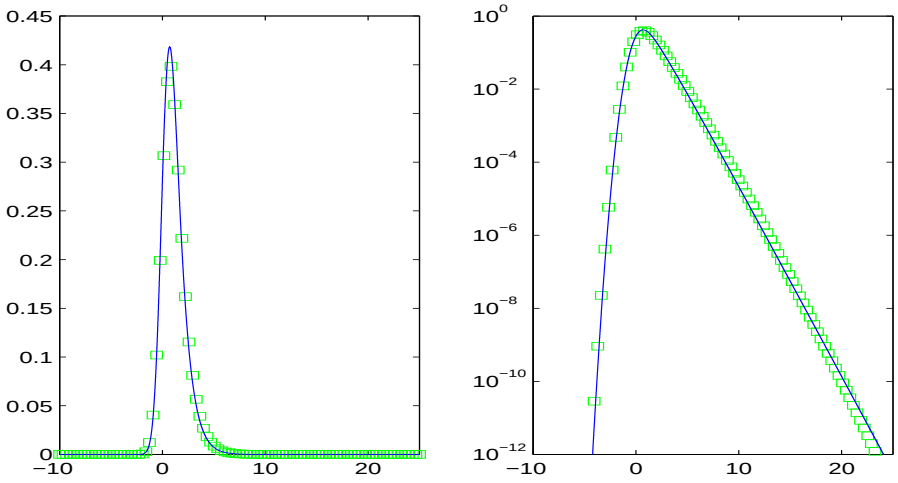


Figure 6.9: Path integration solution (□) and iFFT computed density (-) with parameter setting 1 at $t = 10$. The iFFT computation was performed with $N = 2^{14}$ points.

in table 6.2, and compared with densities computed with iFFT from (A.15). The results for $T = 10$ are presented in figures 6.9, 6.10 and plots of the time evolution of $\mathcal{L}[Y_t]$ are presented in figures 6.11a, 6.11b. Also here we see that there is good accordance between the path integration procedure and the iFFT computed density down to 10^{-8}.

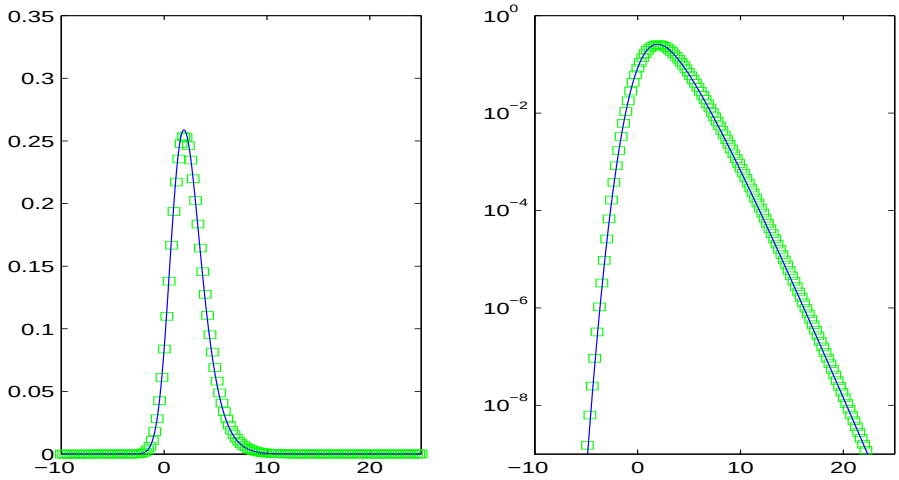
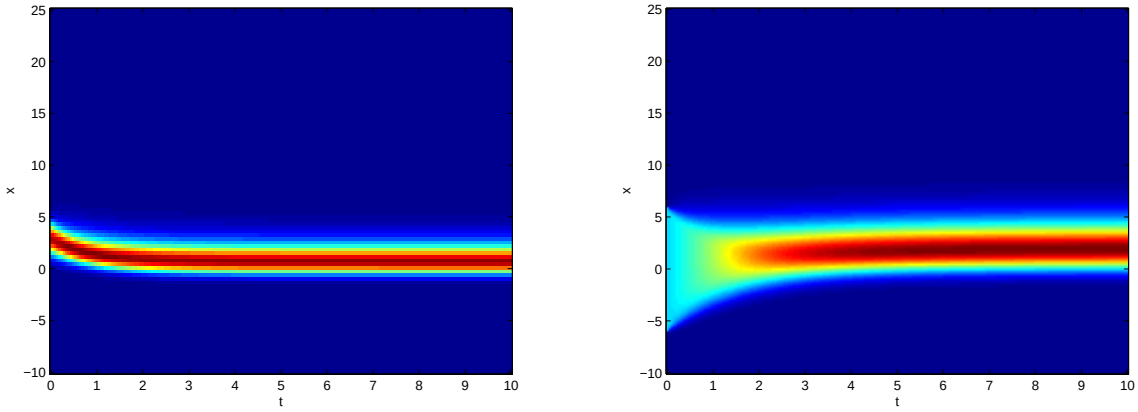


Figure 6.10: Path integration solution (□) and iFFT computed density (-) with parameter setting 2 at $t = 10$. The iFFT computation was performed with $N = 2^{14}$ points.



(a) Parameter Setting 1

(b) Parameter setting 2

Figure 6.11: Illustration of the time-evolution of $\mathcal{L}[Y_t]$.

6.2 Black-Scholes Type Equations

In this section we look at stochastic differential equations on the form

$$\mathrm{d}X_t = aX_t\mathrm{d}t + bX_t\mathrm{d}L_t,\ X_t, a, b \in \mathbb{R},\ t \in [0, T] \tag{6.17}$$

where L_t is the driving Levy process. First we look at the classic Black-Scholes model (Øksendal (2003) or Karatzas & Shreve (1988)) where $L_t = B_t$.

6.2.1 The Classic Black-Scholes Model

This subsection is laid out as follows. First the exact transition kernel is derived. Then we derive the forward Kolmogorov operator, and finally we present some results from the numerical path integration method, using the exact transition kernel as our reference.
Any textbook (e.g. Øksendal (2003)) on Itô stochastic calculus gives the strong solution of (6.17) when $L_t = B_t$. It is given as

$$X_t = X_0 \exp\left(\left(a - \frac{b^2}{2}\right)t + bB_t\right). \tag{6.18}$$

We take $X_0 = x > 0$ to be non-stochastic. It is clear that

$$\mathcal{L}\left[\left(a - \frac{b^2}{2}\right)t + bB_t\right] = N((a - b^2/2)t, b^2t), \tag{6.19}$$

hence the exponential has a Lognormal density (Casella & Berger 2002). Finally using transformation of a stochastic variable with a constant map yields

$$k(y, x, t) = \frac{1}{\sqrt{2\pi b^2 t}y} \exp\left(-\frac{(\log(y/x) - \mu(t))^2}{2b^2t}\right) \tag{6.20}$$

where $\mu(t) = (a - b^2/2)t$.
The derivation of the forward Kolmogorov equation is also straight forward. We use (2.35) to get

$$A^*k(y) = -D_y(ayk(y)) + \frac{1}{2}D_{y,y}(b^2y^2k(y)). \tag{6.21}$$

It is easy to verify that (6.20) solves the forward Kolmogorov equation with operator (6.21).
In the numerical implementations with multiplicative noise, we approximate the path integration integral directly. In this case the time discrete transition kernel is given as

$$\bar{k}(y, x, \tau) = \frac{1}{\sqrt{2\pi(xb)^2\tau}} \exp\left(-\frac{(y - x - ax\tau)^2}{2(bx)^2\tau}\right). \tag{6.22}$$

We run the numerical path integration procedure with parameter settings given in table 6.3 and $T = 1$ in both cases. $Q = 201$ grid points and $N = 2^6$ quadrature points were used. The densities at T are given in figures 6.12, 6.13. We see that the numerical path integration solution is accurate at least down to density levels of order 10^{-8} in both cases.

Setting	a	b	X_0	τ
1	0.3	0.15	12	1/200
2	0.5	0.1	12	1/200

Table 6.3: Parameter settings for the Black Scholes equation driven by Brownian motion.

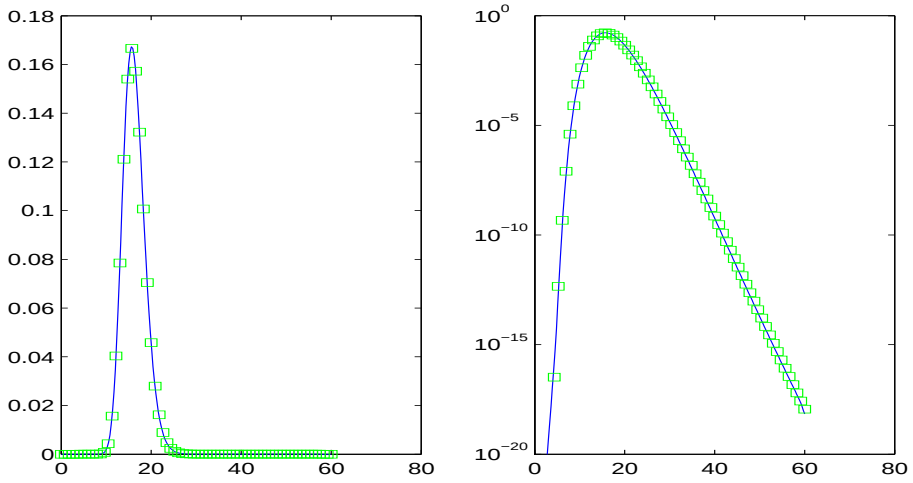


Figure 6.12: Path integration solution (-) and analytical density (□) with parameter setting 1 at $T = 1$.

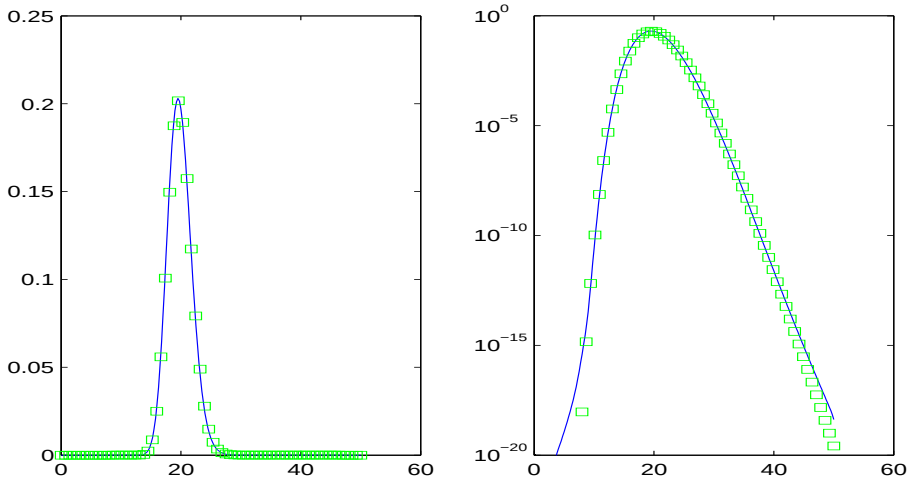


Figure 6.13: Path integration solution (-) and analytical density (□) with parameter setting 1 at $T = 1$.

characteristic	NIG($\delta,\ \alpha,\ \beta$)
mean	$\delta\beta/\sqrt{\alpha^2-\beta^2}$
variance	$\alpha^2\delta(\alpha^2-\beta^2)^{(-3/2)}$
skewness	$3\beta\alpha^{-1}\delta^{-1/2}(\alpha^2-\beta^2)^{(-1/4)}$
kurtosis	$3\left(1+\frac{\alpha^2+4\beta^2}{\delta\alpha^2\sqrt{\alpha^2-\beta^2}}\right)$

Table 6.4: Some characteristics of the NIG density.

6.2.2 The Black-Scholes Equation driven by a Normal Inverse Gaussian Process

In this example we look at the process determined by the stochastic differential equation

$$\mathrm{d}X_t = aX_t\mathrm{d}t + bX_t\mathrm{d}L_t \tag{6.23}$$

where L_t is a normal inverse Gaussian (NIG) process (Barndorff-Nielsen 1998), (Schoutens 2003). The NIG process is a three parameter process commonly used in mathematical finance since heavy tails and skewness can be incorporated. Its characteristic function is given as

$$\phi_{L_t}(u) = \exp\left(-\delta t[\sqrt{\alpha^2-(\beta+iu)^2}-\sqrt{\alpha^2-\beta^2}]\right) \tag{6.24}$$

where $\delta > 0$ can be thought of as a scale-parameter, $\alpha > 0$ determines the tail behavior and $\beta,\ |\beta| < \alpha$ the skewness. To gain a better understanding of the NIG density, some characteristics are listed in table 6.4. In our numerical path integration implementation we need the transition kernel of a process on the form $\hat{L}_t = cL_t$ where $c \in \mathbb{R}$. We use (6.24) to find the corresponding parameters of $\hat{L}_t$:

$$\begin{aligned}\phi_{\hat{L}_t}(u) &= \exp\left(-\delta t[\sqrt{\alpha^2-(\beta+icu)^2}-\sqrt{\alpha^2-\beta^2}]\right)\\ &= \exp\left(-\delta t[\sqrt{c^2\alpha^2/c^2-c^2(\beta/c+iu)^2}-\sqrt{c^2\alpha^2/c^2-c^2\beta^2/c^2}]\right)\\ &= \exp\left(-\delta tc[\sqrt{\alpha^2/c^2-(\beta/c+iu)^2}-\sqrt{\alpha^2/c^2-\beta^2/c^2}]\right).\end{aligned} \tag{6.25}$$

Hence $\hat{L}_t$ has parameters $\hat{\delta} = |c|\delta$, $\hat{\alpha} = \alpha/|c|$ and $\hat{\beta} = \beta/c$. From this is it relatively easy to derive the time-discrete transition kernel suitable for direct integration. The density of the NIG process with parameters (δ, α, β) is given as (Schoutens 2003)

$$d_{L_t}(y) = \frac{\alpha\delta t}{\pi}\exp(\delta t\sqrt{\alpha^2-\beta^2}+\beta y)\frac{K_1(\alpha\sqrt{(\delta t)^2+y^2})}{\sqrt{(\delta t)^2+y^2}} \tag{6.26}$$

Setting	a	b	α	β	δ	$\mathcal{L}[X_0]$	τ
1	0.5	0.1	5.0	0.0	1.5	$N(75, 0.5)$	1/100
2	0.5	0.1	5.0	-1.0	1.5	$N(75, 0.5)$	1/100

Table 6.5: Parameter settings for the normal inverse Gaussian Black-Scholes model.

where K_1 is the modified Bessel function of third kind with parameter 1 (Schoutens 2003). From (6.26) it is easy to see that the time-discrete transition kernel is given as

$$\bar{k}(y,x,\tau) = \frac{\alpha\delta\tau}{\pi}\exp\left(\delta\tau\sqrt{\alpha^2-\beta^2}+\frac{\beta(y-(1+a\tau)x)}{bx}\right) \times \frac{K_1\left(\frac{\alpha}{bx}\sqrt{(\delta\tau bx)^2+(y-(1+a\tau)x)^2}\right)}{\sqrt{(\delta\tau bx)^2+(y-(1+a\tau)x)^2}}. \quad (6.27)$$

The numerical path integration method is implemented using direct integration. In fact we use the same code as for the previous Black-Scholes equation, only altering the transition kernel (6.27). The modified Bessel function K_1 is computed using `tbessk.f90` (Moreau 2005). The code is tested using the parameter settings given in table 6.5 up to time $T = 1$. As reference we use $1.0e8$ Monte Carlo simulations using the Euler scheme with 500 time-steps using `MATLAB` (Higham 2001). The random increments are generated using `randraw.m` (Bar-Guy 2005). We use kernel density estimation with Gaussian kernels and parameter $h = 0.5$ (Venables & Ripley 2003). Since the accuracy of the kernel density estimate decays as fewer individual paths end in the tails of the density, we only compute the kernel density estimate on the interval were 1000 paths end outside on each side.
The results are presented in figures 6.14, 6.15 and images of the time-evolutions of the densities are given in figures 6.16a, 6.16b. As seen in the figures, the numerical path integration method seems to give accurate results in the high density region. What happens in the low density regions is at this point hard to tell, since we need wast amounts of Monte Carlo simulations to get accurate kernel density estimates. Figures 6.16a and 6.16b give a nice quantitative image about what happens when the driving noise is not a martingale. We see that for parameter setting 2, the left-skewed transition kernel approximately balances the positive drift.

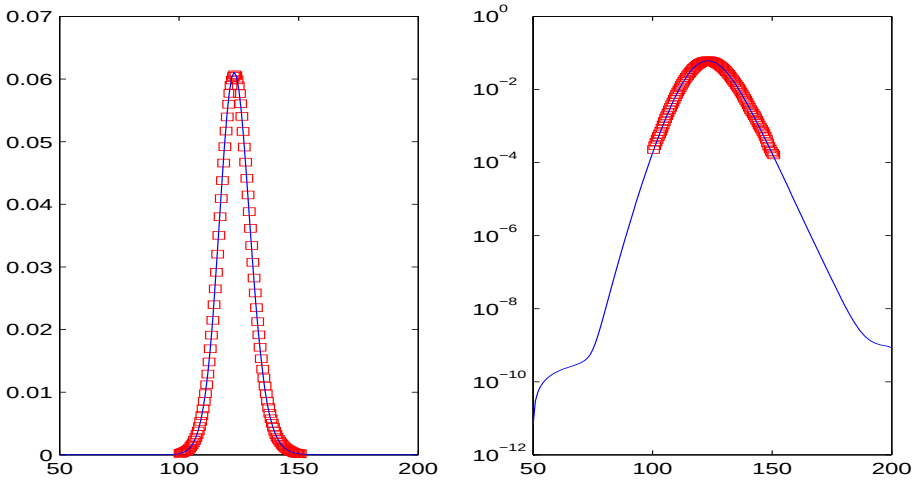


Figure 6.14: Numerical path integration solution (-) and kernel density estimate based on Monte Carlo simulation (□).

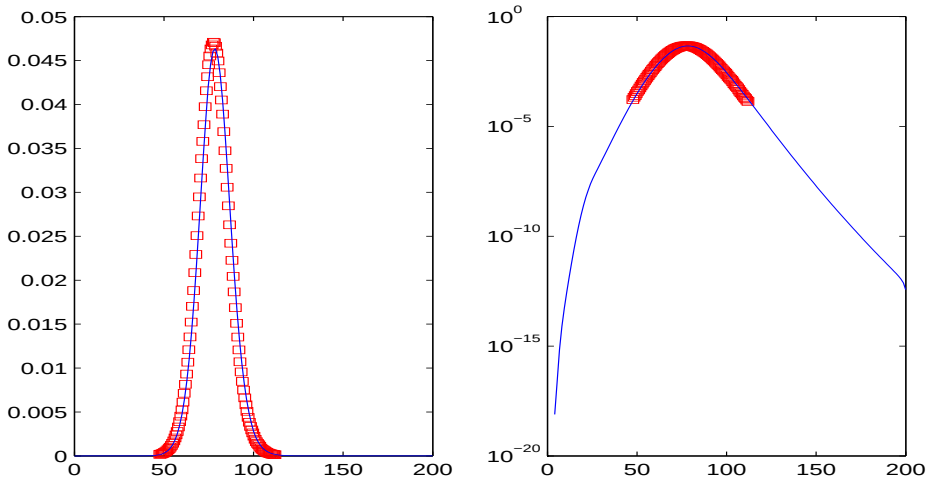
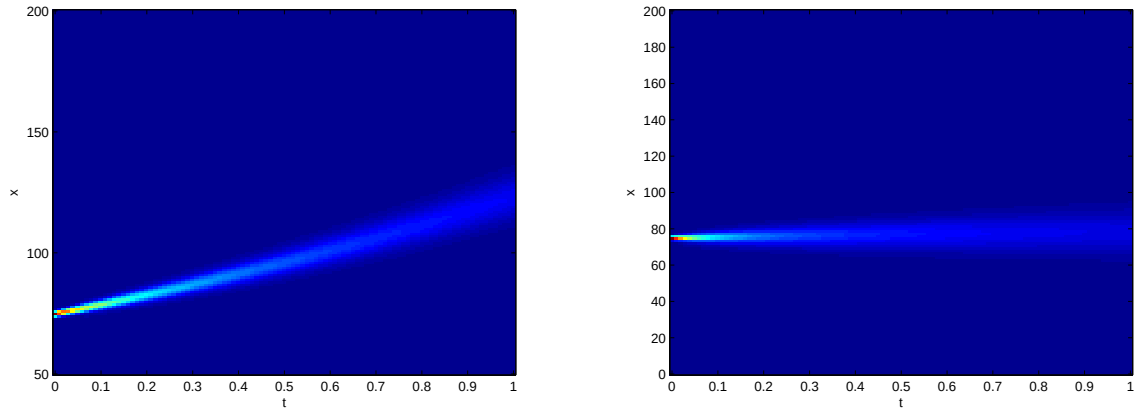


Figure 6.15: Numerical path integration solution (-) and kernel density estimate based on Monte Carlo simulation (□).



(a) Parameter Setting 1

(b) Parameter setting 2

Figure 6.16: Illustration of the time-evolution of $\mathcal{L}[X_t]$ where X_t is given in (6.23).

6.3 A Nonlinear Equation

Our last example is taken from Kloeden & Platen (1999). It can be shown that the Itô stochastic differential equation

$$\mathrm{d}X_t = -a^2 X_t(1 - X_t^2)\mathrm{d}t + a(1 - X_t^2)\mathrm{d}B_t, \qquad X_t, a \in \mathbb{R},\ t \in [0, T] \quad (6.28)$$

has strong solution

$$X_t = \tanh(aB_t + \operatorname{arctanh}(X_0)). \quad (6.29)$$

Since tanh is monotonous, it is easy to derive the transition kernel of the exact solution. Take $X_0 = x,\ -1 < x < 1$ to be non-stochastic. Then $\mathcal{L}[aB_t + \operatorname{arctanh}(x)] = N(\operatorname{arctanh}(x), a^2 t)$, and using that X_t is a mapped Gaussian variable we obtain (Casella & Berger 2002)

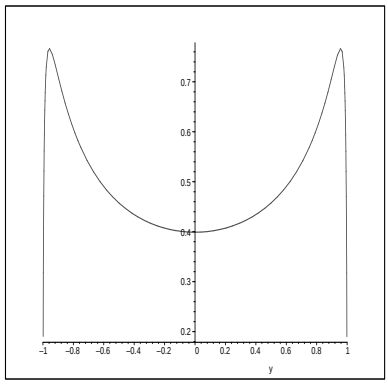


Figure 6.17: Plot of the transition kernel (6.30).

$$k(y, x, t) = \frac{1}{\sqrt{2\pi a^2 t}(1 - y^2)} \exp\left(-\frac{(\operatorname{arctanh}(y) - \operatorname{arctanh}(x))^2}{2a^2 t}\right),\ y \in [-1, 1]. \quad (6.30)$$

It can be shown that

$$\lim_{|y| \nearrow 1} k(y, x, t) = 0 \ \forall\ x \in (-1, 1),\ t < \infty \quad (6.31)$$

hence the forward transition kernel has compact support. The forward transition kernel $k(y, 0, 1)$ is plotted in figure 6.17 for $a = 1$. From figure 6.17 we see that the exact solution

Setting	a	x	τ
1	0.1	0	3/200
2	0.5	0	3/200
3	0.5	1/2	3/200

Table 6.6: Parameter settings for (6.28)

has two modes, which becomes increasingly sharp as t grows, making this problem harder numerically.
The Forward Kolmogorov operator is easy to find given (2.35), i.e.

$$A^*(y) = -D(-a^2 y(1-y^2)k(y)) + \frac{1}{2}D^2(a^2(1-y^2)^2 k(y)). \tag{6.32}$$

Also here it is easy to see that the transition kernel k satisfies the Forward Kolmogorov equation corresponding to (6.32).
The numerical path integration method was applied to (6.28). The time discrete transition kernel is given as

$$\bar{k}(y,x,\tau) = \frac{1}{\sqrt{2\pi\tau}a(1-x^2)} \exp\left(-\frac{(y-x+a^2x(1-x^2))^2\tau}{2(a(1-x^2))^2\tau}\right) \tag{6.33}$$

and we used direct integration of the path integration integral. We do three runs with parameters given in table 6.6. We set $T = 3$ and run the simulation with relatively short time-steps, required by the nonlinearity of the model. We use $Q = 151$ grid-points and $N = 200$ quadrature points. Numerical path integration densities and exact densities are plotted in figures 6.18, 6.19 and 6.20. The final density from a run with parameter setting 3 up to $T = 10$ is presented in figure 6.21. Images illustrating the time-evolution of the densities are given in figures 6.22, 6.23 and 6.24. The code is not altered to make the compact support embedded, and the computations are done on the domain $[-1.2, 1.2]$. We see that the numerical path integration code performs well in the domains where there is moderately high probability density when the exact density is relatively smooth. As shown in figure 6.21, the numerical path integration method has trouble following the dynamics when the densities are close to non-smooth. To cope with this, finer spatial grids have to be considered.

6.4 Concluding Remarks

As we have seen in the examples above, the numerical path integration method gives good results on a quite wide range of 1-dimensional problems. It is relatively easy to write a path integration code for such problems, all the codes used in the examples consist of approximately 150 lines, excluding spline-libraries, iFFTs and so on.
There are two other main methods for solving such problems - Monte Carlo simulations

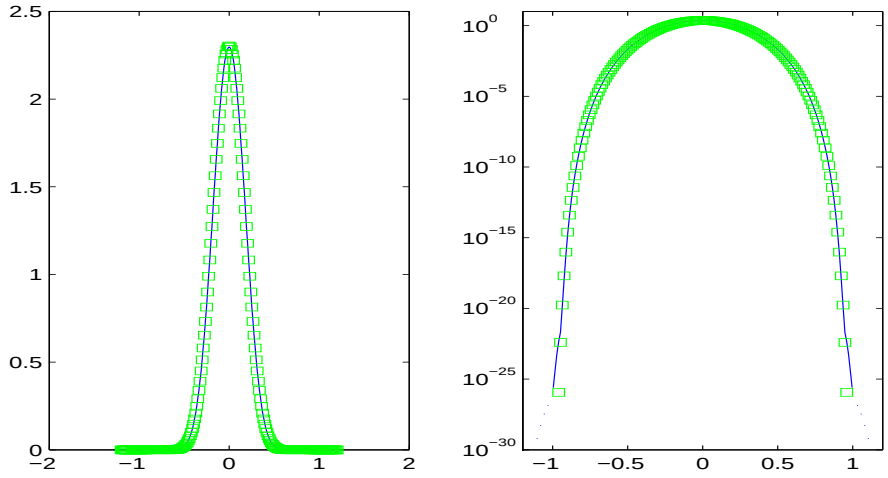


Figure 6.18: Numerical path integration solution (-) and analytical density (□) with parameter setting 1 at $T = 3$.

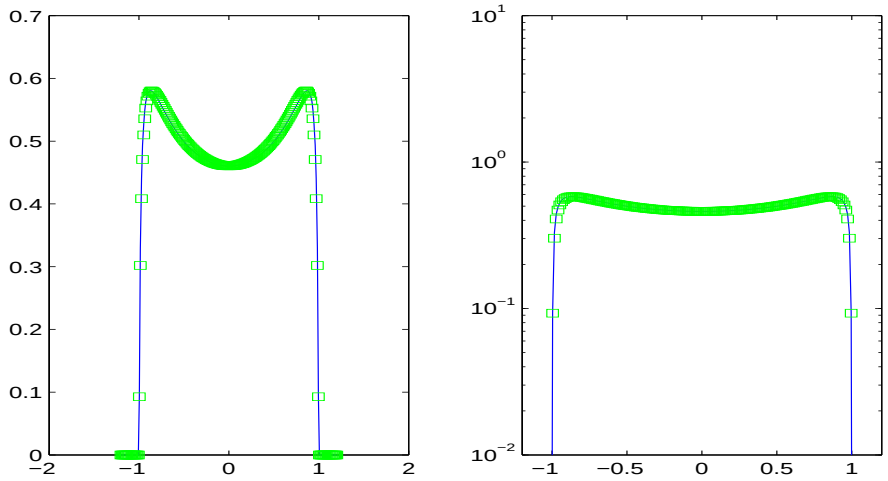


Figure 6.19: Numerical path integration solution (-) and analytical density (□) with parameter setting 2 at $T = 3$.

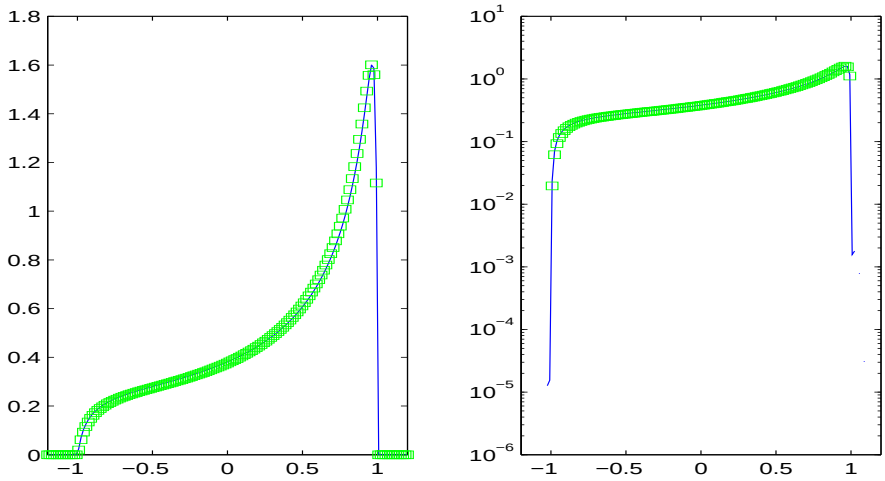


Figure 6.20: Numerical path integration solution (-) and analytical density (□) with parameter setting 3 at $T = 3$.

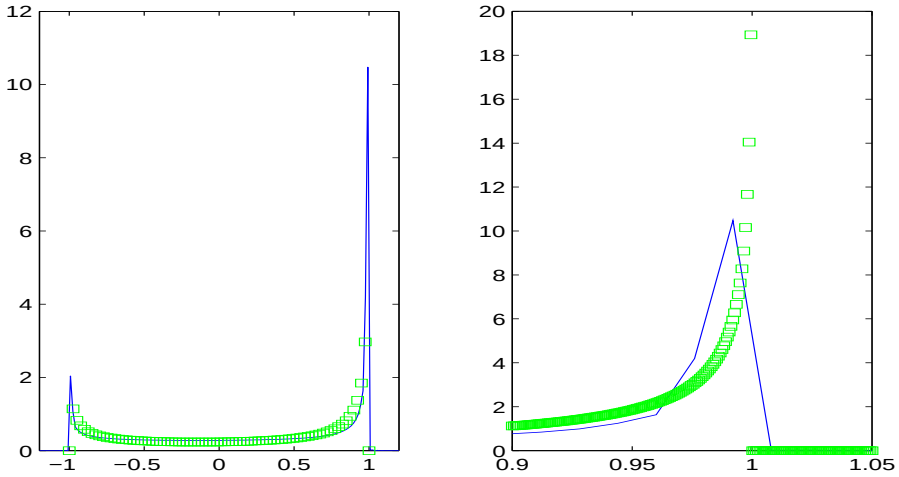


Figure 6.21: Numerical path integration solution (-) and analytical density (□) with parameter setting 3 at $T = 10$. The right plot is just magnification of the left. Notice that at this time, the sharp modes of the analytical density makes it hard for the numerical path integration method to follow.

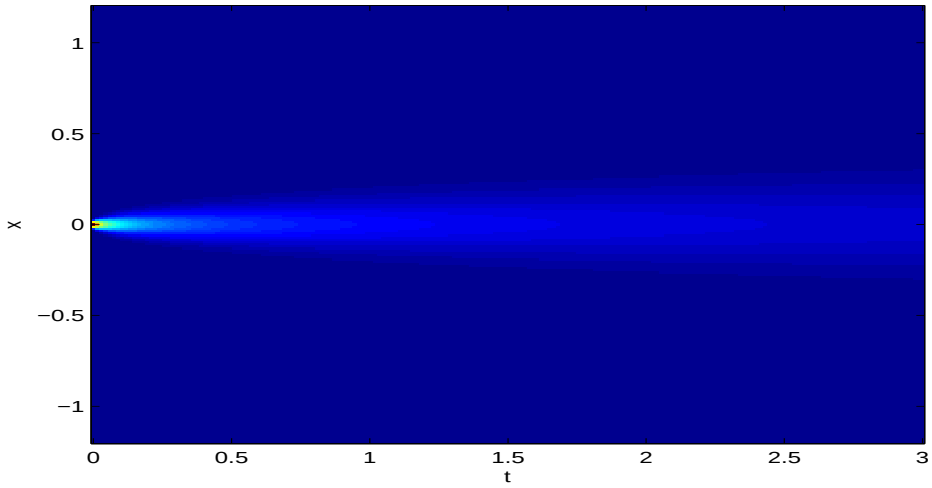


Figure 6.22: Illustration of the time evolution of the numerical path integration density with parameter setting 1.

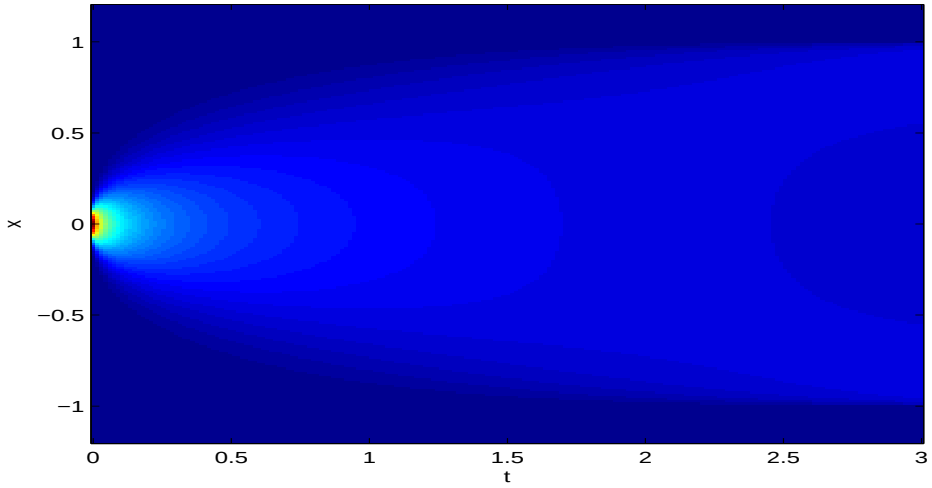


Figure 6.23: Illustration of the time evolution of the numerical path integration density with parameter setting 2. Notice that we see form the strong solution that changing a is equivalent to changing the time scale. Hence this figure 6.22 is just the first part of this picture.

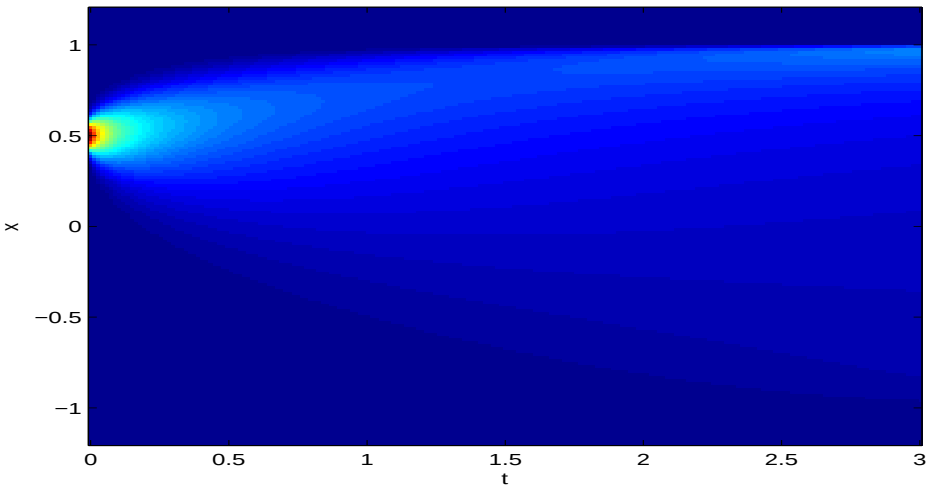


Figure 6.24: Illustration of the time evolution of the numerical path integration density with parameter setting 3.

(see e.g. Kloeden & Platen (1999) or Protter & Talay (1997)) and various kinds of finite difference and finite element methods (see e.g. Briani, La Chioma & Natalini (2004) or Cont & Voltchkova (2005)). The numerical path integration method is superior to Monte Carlo methods in many cases since the CPU-times needed to get desired accuracy is larger. In the example with the Black-Scholes model driven by a NIG-process, the time consumption is $\sim 1-10$ seconds, whereas the Monte Carlo simulation takes approximately half an hour. The downside of the path integration method compared to Monte Carlo methods is that the path integration method is somewhat more complicated to implement, and has more parameters that needs to be tuned for best performance.
Little comparison with the finite difference and finite element methods has been done. What seems clear is that the CPU-times are comparable with those of the path integration method. Upsides for the path integration method is that the codes can be reused when the model changes, whereas the model generally is more embedded in finite difference methods. A downside comparing with methods from the integro-PDE tool box is that more explicit convergence rates are at hand.

Chapter 7

Conclusions

7.1 Conclusions

At the end of each previous chapter, we have given preliminary concluding remarks. This conclusion is therefore relatively brief. What this thesis shows, is that the numerical path integration method can, with relative ease, be extended to Lévy driven stochastic differential equations. If a closed form expression for the time-discrete transition kernel exits, the extension is straight forward. In cases where the Lévy process is specified by a Lévy triplet, iFFT-methods can be applied to approximate the transition kernels. However for models with multiplicative noise, this might amount to computing a vast amount of densities, making the path integration method computationally inefficient.

7.2 Further Work

To finish of, we mention a few direction for further work:

1. Extending the numerical path integration method for Lévy driven stochastic differential equations to higher dimensions. This is essential in many applications since the models have an embedded n-dimensional nature.
2. Comparing the numerical path integration method with finite difference schemes for integro-PDEs.
3. Using the numerical path integration method as basis for parameter estimation using maximum-likelihood. Automatic Differentiation might be applied to compute gradients of the maximum-likelihood function.
4. Extending path integration in Fourier space (see appendix A) to nonlinear equations using FFT-methods.

Bibliography

Applebaum, D. (2004), *Levy Processes and Stochastic Calculus*, Cambridge Studies in Advanced Mathematics, Cambridge University Press.

Bally, V. & Talay, D. (1995), 'The law of the Euler Scheme for Sthochastic Differential Equations (i) : convergence rate of the distribution function', *Probability Theory and Related Fields* **104**, 43–60.

Bally, V. & Talay, D. (1996), 'The law of the Euler Scheme for Stochastic Differential Equations (ii) : Convergence Rate of the Density', *Monte Carlo Methods and Applications* **2**, 93–128.

Bar-Guy, A. (2005), *randraw.m*, `http://www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=7309`.

Barndorff-Nielsen, O. E. (1998), 'Processes of normal inverse Gaussian type', *Finance Stoch.* **2**(1), 41–68.

Bass, R. F. (2004), 'Stochastic differential equations with jumps', *Probability Surveys* **1**, 1–19.

Briani, M., La Chioma, C. & Natalini, R. (2004), 'Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory', *Numer. Math.* **98**(4), 607–646.

Burden, R. L. & Faires, J. D. (2000), *Numerical Analysis*, 7th edn, Brooks Cole.

Burkardt, J. (1999), *spline.f90*, `http://www.scs.fsu.edu/~burkardt/f_src/spline/spline.html`.

Casella, G. & Berger, R. L. (2002), *Statistical Inference, Second Edition*, Duxbury/Thompson Learning.

Champeney, D. C. (1987), *A handbook of Fourier Theorems*, Cambridge University Press.

Cont, R. & Tankov, P. (2004), *Financial Modelling with Jump Processes*, Chapman & Hall/CRC.

Cont, R. & Voltchkova, E. (2005), 'A finite difference scheme for option pricing in jump diffusion and exponential Lévy models', *SIAM J. Numer. Anal.* **43**(4), 1596–1626 (electronic).

Daly, E. & Porporato, A. (2006), 'Probabilistic dynamics of some jump-diffusion systems', *Physical Review E* **73**.

deBoor, C. (2001), *A Practical Guide to Splines*, number 27 *in* 'Applied Mathematical Sciences', Springer Verlag.

Evans, L. C. (1998), *Partial Differential Equations*, American Mathematical Society.

Feng, G. M., Wang, B. & Lu, Y. F. (1992), 'Path Integral, Functional Method, and Stochastic Dynamical Systems', *Probabilistic Engineering Mechanics* **7**.

Folland, G. B. (1999), *Real Analysis : Modern Techniques and Their Applications*, Pure and Applied Mathematics, 2 edn, Wiley-Interscience.

Gasquet, C. & Witomski, P. (1999), *Fourier Analysis and Applications*, Springer-Verlag.

Gihman, I. I. & Skorohod, A. V. (1972), *Stochastic Differential Equations*, Springer-Verlag.

Hanson, F. B. (2006), *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis and Computation*, SIAM. The text used is the draft available on "`http://www.math.uic.edu/~hanson/math574/#Text`" as of July 8, 2005. The book is to be published in 2006.

Higham, D. J. (2001), 'An algorithmic introduction to numerical simulation of stochastic differential equations', *SIAM Review* **43**(3), 525–546.

Higham, D. J., Mao, X. & Stuart, A. M. (2002), 'Strong Convergence of Euler-type Methods for Nonlinear Stochastic Differential Equations', *SIAM Journal of Numerical Analysis* **40**(3).

Hughett, P. (1998), 'Error Bounds for Numerical Inversion of a Probability Characteristic Function', *SIAM Journal of Numerical Analysis* **35**(4).

Jacod, J., Kurtz, T. G., Mèlèard, S. & Protter, P. (2005), 'The Approximate Euler Method for Levy Driven Sthochastic Differential Equations', *Annales de l'Institut Henri Poincaré* **41**, 523–558.

Karatzas, I. & Shreve, S. (1988), *Brownian Motion and Stochastic Calculus*, Springer Verlag.

Kloeden, P. E. & Platen, E. (1999), *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag.

Kolnes, F. E. (2004), The path integration method with applications to gear dynamics, Master's thesis, Department of Mathematical Sciences - NTNU.

Lasota, A. & Mackey, M. (1994), *Chaos, Fractals, Noise*, Springer-Verlag.

Milstein, G. (1995), *Numerical Integration of Stochastic Differential Equations*, Kluwer Academic Publishers.

Milstein, G. N., Platen, E. & Schurz, H. (1998), 'Balanced Implicit Methods for Stiff Stochastic Systems', *SIAM Journal on Numerical Analysis* **35**(3).

Moreau, J.-P. (2005), *tbessk.f90*, `http://perso.wanadoo.fr/jean-pierre.moreau/Fortran/tbessk_f90.txt`.

Naess, A. (2001), 'Lecture Notes on the Numerical Solution of Stochastic Differential Equations', *Preprint Statistics - NTNU* **2001**(11).

Naess, A. & Moe, V. (2000), 'Efficient Path Integration Methods for Nonlinear Dynamical Systems', *Probabilistic Engineering Mechanics* **15**.

Onsager, L. & Machlup, S. (1953), 'Fluctuations and Irreversible Processes', *Physical Reveiw* **91**(8).

Protter, P. E. (2004), *Stochastic Integration and Differential Equations*, second edn, Springer Verlag.

Protter, P. & Talay, D. (1997), 'The Euler Scheme for Lévy Driven Stochastic Differential Equations', *The Annals of Probability* **25**(1).

Risken, H. (1984), *The Fokker-Planck Equation*, Springer-Verlag.

Rudin, W. (1976), *Principles of Mathematical Analysis, Third ed.*, McGraw-Hill.

Sato, K. (1999), *Levy Processes and Infinitely Divisible Distributions*, Cambridge Studies in Advanced Mathematics, Cambridge University Press.

Schoutens, W. (2003), *Lévy Processes in Finance - Pricing Financial Derivatives*, Wileys Series in Probability and Statistics, Wiley.

Schultz, M. H. (1973), *Spline Analysis*, Series in Automatic Computation, Prentice-Hall.

Stroock, D. (1998), *A Concise Introduction to the Theory of Integration*, 3 edn, Birkhauser.

Sun, J. Q. & Hsu, C. S. (1990), 'The Generalized Cell Mapping Method in Nonlinear Randon Vibration Based Upon Short-Time Gaussian Approximation', *Transactions of the ASME* **57**.

Sun Studio 11: Sun Performance Library Reference Manual - ZFFTD (2005), `http://docs.sun.com/source/819-3691/zfftd.html`.

Troutman, J. L. (1996), *Variational Calculus and Optimal Control - Optimization with Elementary Convexity*, second edn, Springer Verlag.

Venables, W. & Ripley, B. (2003), *Modern Applied Statistics with S*, 3rd edn, Springer Verlag.

Wissel, C. (1979), 'Manifolds of Eqvivalent Path Integral Solutions of the Fokker-Planck Equation', *Zeitschrift fur Physik B* **35**.

Yan, L. (2005), 'Asymptotic Error for the Milstein Scheme for SDEs Driven by Continuous Semimartingales', *The Annals of Applied Probability* **15**(4), 2706–2738.

Øksendal, B. (2003), *Stochastic Differential Equations, 6. Edition*, Springer Verlag.

Øksendal, B. & Sulem, A. (2005), *Applied Stochastic Control of Jump Diffusions*, Springer Verlag.

Appendix A

Path Integration in Fourier Space

A.1 Path Integration in Fourier Space for the Langevin Equation

This section is a digression from the path integration exposition in this text. In the case when the stochastic differential equation is linear with additive noise, i.e. the Langevin equation, it is simple to derive the characteristic function of the law of the time-discrete solution. Letting $\tau \to 0$ we can in many cases obtain the characteristic function of the exact solution as well.
We use the framework

$$\mathrm{d}X_t = -aX_t\mathrm{d}t + b\mathrm{d}L_t,\ X_t \in \mathbb{R},\ t \in [0,T],\ a,b > 0 \tag{A.1}$$

where L_t is some Levy process with characteristic function $\phi_{L_t}(u)$ and X_0 is a stochastic variable with characteristic function $\psi_{X_0}(u)$. Then we have the following result:

Result A.1: *The characteristic function of the n-th step in the Euler scheme applied on (A.1) is given as:*

$$\psi_{\bar{X}_n}(u) = \psi_{X_0}(c^n u)\prod_{j=0}^{n-1}\phi_{L_\tau}(c^j bu) \tag{A.2}$$

where $c = 1 - a\tau$.

Proof. We start with the time-discrete Euler scheme with time discretization $\tau = T/K,\ t_i = i\tau,\ i = 0,1,\ldots,K$:

$$\bar{X}_{i+1} = \bar{X}_i - a\tau\bar{X}_i + b[L_{t_{i+1}} - L_{t_i}] = (1-a\tau)\bar{X}_i + b[L_{t_{i+1}} - L_{t_i}]. \tag{A.3}$$

We assume K to be large enough such that $|1 - a\tau| < 1$. The proof is of induction type. It is clear from (A.3) that

$$\psi_{\bar{X}_{i+1}}(u) = \psi_{\bar{X}_i}(cu)\phi_{L_\tau}(bu). \tag{A.4}$$

hence the for $i = 1$ the induction hypothesis holds. Assume the induction hypothesis holds for $n = k$. Then

$$\begin{aligned}\psi_{\bar{X}_{k+1}}(u) &= \psi_{\bar{X}_k}(cu)\phi_{L_\tau}(bu)\\ &= \left[\psi_{X_0}(c[c^k u])\prod_{j=0}^{k-1}\phi_{L_\tau}(c[c^j bu])\right]\phi_{L_\tau}(bu) = \psi_{X_0}(c^{k+1}u)\prod_{j=0}^{k}\phi_{L_\tau}(c^j bu).\end{aligned} \tag{A.5}$$

□

We apply this result in some examples:

Example A.1: We can formally derive the exact transition kernel of the Brownian motion driven Langevin equation using (A.2): We set $X_0 = x$, i.e. non stochastic and $\psi_{X_0}(u) = \exp(iux)$. Further, the characteristic function of a Brownian increment is given as $\phi_{B_\tau}(u) = \exp(-\tau u^2/2)$. Hence

$$\begin{aligned}\psi_{\bar{X}_M}(u) &= \exp(iuxc^M) \prod_{j=0}^{M-1} \exp(-\tau(bu)^2 c^{2j}/2) \\ &= \exp\left(iuxc^M - \frac{1}{2}\tau(bu)^2 \sum_{j=0}^{M-1} (c^2)^j \right) = \exp\left(iuxc^M - \frac{1}{2}\tau(bu)^2 \frac{1-c^{2M}}{1-c^2} \right) \end{aligned} \quad (A.6)$$

We now substitute $\tau = T/M$ and $c = 1 - aT/M$:

$$\begin{aligned}\psi_{\bar{X}_M}(u) &= \exp\left(iux(1-\frac{aT}{M})^M - (bu)^2 \frac{T}{2M} \frac{1-(1-\frac{aT}{M})^{2M}}{1-(1-\frac{aT}{M})^2} \right) \\ &= \exp\left(iux(1-\frac{aT}{M})^M - (bu)^2 \frac{1-(1-\frac{aT}{M})^{2M}}{2(2a-\frac{T}{M})} \right) \end{aligned} \quad (A.7)$$

Now we use that (Rudin 1976)

$$\lim_{M\to\infty} \left(1 - \frac{k}{M}\right)^M = \exp(-k), \quad (A.8)$$

hence by letting $M \to \infty$ we obtain

$$\psi_{X_T}(u) = \exp\left(iu[x\exp(-aT)] - \frac{u^2}{2}\left[\frac{b^2}{2a}(1-\exp(-2aT))\right] \right). \quad (A.9)$$

Clearly this is accordance with the results in example 5.1

Our next example is the Langevin equation driven by Brownian motion and a compound Poisson process with zero-mean Gaussian jumps. We derived the characteristic function for an increment of this process in subsection 4.2. The results from this example are used in subsection 6.1.2.

Example A.2: Recall from subsection 4.2 that the characteristic function of the process consisting of Brownian motion and a compound Poisson process with Gaussian jumps is given as

$$\phi_{U_\tau}(u) = \exp\left(\tau(\frac{u^2}{2} + \lambda(\exp(-\frac{\sigma_Z^2 u^2}{2}) - 1) \right) \quad (A.10)$$

We take the initial condition to be deterministic, that is $X_0 = x$. Using the derivation above for the drift and Brownian part, we get that

$$\psi_{\bar{X}_M}(u) = \exp\left(iux(1-\frac{aT}{M})^M - \frac{(bu)^2}{2(2a-T/M)}(1-(1-\frac{aT}{M})^{2M}) \right.$$
$$\left. + \lambda\frac{T}{M}\sum_{j=0}^{M-1}\left[\exp\left(-\frac{(b\sigma_Z u)^2}{2}(1-\frac{aT}{N})^{2j}\right)-1\right]\right) \quad \text{(A.11)}$$

Letting $M \to \infty$ we obtain

$$\psi_{X_T}(u) = \exp\left(iux\exp(-aT) - \frac{(bu)^2}{4a}(1-\exp(-2aT)) \right.$$
$$\left. + \frac{\sqrt{2\pi}\lambda b\sigma_Z u}{8a}\left[-\mathsf{erf}\left(\frac{b\sigma_Z u}{\sqrt{2}}\right)+\exp(-aT)\mathsf{erf}\left(\frac{b\sigma_Z u}{\sqrt{2}}\exp(-aT)\right)\right]\right) \quad \text{(A.12)}$$

Performing an inverse Fourier transform on this expression seem hard, but we can use the methods described in chapter 4 to approximate the exact transition kernel.

Finally we derive the characteristic function of the solution Langevin equation driven by the asymmetrical compound Poisson process described in subsection 4.3.

Example A.3: Recall that the characteristic function of the process consisting of Brownian motion and a compound Poisson process with exponentially distributed jumps is given as

$$\phi_{U_t} = \exp\left(-\frac{\tau}{2}u^2 + \lambda\tau\left(\frac{1}{1-iu\beta}-1\right)\right) \quad \text{(A.13)}$$

Again we assume non-stochastic initial condition x. Using the derivation in example A.1 we have that

$$\psi_{\bar{X}_M}(u) = \exp\Big(iux(1-\frac{aT}{M})^M - \frac{(bu)^2}{2(2a-T/M)}(1-(1-\frac{aT}{M})^{2M}) \Big.$$
$$\Big. + \lambda\frac{T}{M}\sum_{j=0}^{M-1}\frac{iub\beta(1-\frac{aT}{M})^j}{1-iub\beta(1-\frac{aT}{M})^j}\Big). \quad \text{(A.14)}$$

Letting $M \to \infty$ we obtain the Fourier transform of the transition kernel for the exact solution:

$$\psi_{X_T}(u) = \exp\Big(iux\exp(-aT) - \frac{(bu)^2}{4a}(1-\exp(-2aT)) \Big.$$
$$\Big. + \frac{\lambda}{a}(\log(1-iub\beta\exp(-aT)) - \log(1-iub\beta))\Big) \quad \text{(A.15)}$$

We use this as our reference for the numerical path integration solution presented in subsection 6.1.3.

Using (A.8), it is in many cases relatively simple to introduce stochastic initial conditions. E.g. Gaussian initial conditions lead to terms of form $-(\sigma_0 u^2)\exp(-2aT)/2$ in the exponent.
In a numerical setting, (A.2) might form a basis for a numerical method for small τ. This is not explored here due to time constraints on this work.

A.2 Path Integration in Fourier Space for Nonlinear Additive Noise Equations

In this section we propose a method for path integration in Fourier space with f_1 being nonlinear and $f_2(x) = b$. The idea is basically the same as for the Langevin equation, but analytical expressions seem harder to come by. First we assume that $x + \tau f_1(x)$ is monotone.

Result A.2: *The characteristic function of $\bar{X}_{i+1}$ can be computed recursively as*

$$\psi_{\bar{X}_{i+1}}(u) = \zeta_i(u)\phi_{bL_\tau}(u) \tag{A.16}$$

where ζ_i is computed as follows:

$$g = (x + \tau f_1(x))^{-1} \quad \textit{i.e. the inverse} \tag{A.17a}$$
$$d_i = \mathscr{F}^{-1}\psi_{\hat{X}_i} \tag{A.17b}$$
$$\mathcal{M}_i(x) = |Jg(x)|d_i(g(x)) \tag{A.17c}$$
$$\zeta_i = \mathscr{F}\mathcal{M}_i \tag{A.17d}$$

where J denotes the Jacobian.

Proof. The relation is obvious when looking at an Euler step

$$\bar{X}_{i+1} = \bar{X}_i + \tau f_1(\bar{X}_i) + b[L_{t_{i+1}} - L_{t_i}] \tag{A.18}$$

where ζ_i is the characteristic function of the mapped variable $\bar{X}_i + \tau f_1(\bar{X}_i)$ (see e.g. Casella & Berger (2002)). □

This result might form the backbone of a fast numerical solver using the FFT extensively. It is believed that the computationally expensive spline interpolation and evaluation can be avoided, using rather piecewise linear or piecewise constant interpolants. This is due to the fact that we can use very fine spatial grids. Moreover the inverse mapping g and the Jacobian J are only computed once, admitting fast evaluations of the mapped density $\mathcal{M}$, given d. The computational cost is that of two FFTs and a few element-wise vector multiplications for each time-step.
The monotonicity of $x + \tau f_1(x)$ is not absolutely necessary for applying this method. Casella & Berger (2002) give an explicit formula for the mapped density $\mathcal{M}$ in the case

that f_1 is only piecewise monotone.
This method is also applicable in the n-dimensional case. An important feature of the path integration in Fourier space method is then that it is easy to incorporate degenerate transition kernels arising in models with noise terms in only some of the scalar equations. This is often the case in mechanical systems - see e.g. Kolnes (2004). To see the simplifications, assume that $L_t \in \mathbb{R}^m$ and $X_t \in \mathbb{R}^n$. Letting the models be of additive noise type, the coefficient-function $f_2 = B$ is a matrix, i.e. $B \in \mathbb{R}^{n \times m}$. From the definition of the characteristic function we get

$$\phi_{BL_\tau}(u) = \mathbb{E}[\exp(iu^\top (BL_\tau))] = \mathbb{E}[\exp(i(u^\top B)L_\tau)], \quad u \in \mathbb{R}^n. \tag{A.19}$$

This shows that zero-rows in B correspond to setting the corresponding element in u to 0. Hence zero-rows simplifies our computations, rather than making them harder to implement as is the case when we path integrate in state space.
This method has not been implemented due to time constraints, and should be viewed upon as a proposal for further work.

Appendix B

Computer Codes

B.1 File `PDFs.f90`

These codes were used to compute probability densities. The subroutine `iFFTpdfinit` is the one (in slightly altered forms) used to compute all the iFFT computed densities presented in this text.

```
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! Routines for computing probability densities
! Written by Tore S. Kleppe (toresell@stud.math.ntnu.no)
! spring semester 2006
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

! normal pdf
subroutine normalPDF (n,xx,yy,mu,sigmaS)
   implicit none
   integer, parameter:: k8 = selected_real_kind(15,150)
   real(kind=k8), parameter:: pi = 3.1415926535897932_k8
   integer, intent(in):: n
   real(kind=k8),intent(in):: xx(n)
   real(kind=k8),intent(out):: yy(n)
   real(kind=k8),intent(in):: mu
   real(kind=k8),intent(in):: sigmaS
   yy = 1/sqrt(2*pi*sigmaS)*exp(-(mu - xx)**2/(2*sigmaS))
 end subroutine normalPDF

! uniform pdf
subroutine uniformPDF(n,xx,yy,min,max)
implicit none
   integer, parameter:: k8 = selected_real_kind(15,150)
   real(kind=k8), parameter:: pi = 3.1415926535897932_k8
   integer, intent(in):: n
   real(kind=k8),intent(in):: xx(n)
   real(kind=k8),intent(out):: yy(n)
   real(kind=k8),intent(in):: min
   real(kind=k8),intent(in):: max
   integer i
```

```
    real(kind=k8) dens
    dens = (1.0_k8)/(max-min)
    do i=1,n
       if((xx(i) >= min) .and. (xx(i) < max)) then
          yy(i) = dens
       else
          yy(i) = 0.0_k8
       end if
    end do
end subroutine uniformPDF

! iFFT computed pdfs
! Slight changes of this code must be done to fit other pdfs
! the code given here is the one used for the symmetric jump-
! diffusion increment.

subroutine iFFTpdfinit(N,xx,dens,lambda,sigmaZ,b,t,stdvs)
  implicit none
    integer, parameter:: k8 = selected_real_kind(15,150)
    real(kind=k8),parameter:: h =  0.001_k8
    real(kind=k8),parameter:: pi = 3.1415926535897932
    complex(kind=k8),parameter:: im = (0.0_k8,1.0_k8)

    integer,intent(in):: N ! number of fft points
    real(kind=k8),intent(out):: xx(N) ! abcissa
    real(kind=k8),intent(out):: dens(N) ! density
    real(kind=k8),intent(in):: lambda ! parameters
    real(kind=k8),intent(in):: sigmaZ ! parameters
    real(kind=k8),intent(in):: b ! parameters
    real(kind=k8),intent(in):: t ! parameters
    real(kind=k8),intent(in):: stdvs ! number of standard deviations

    integer Nhalf,NhalfOne
    integer i,ifac(128),ierr
    real(kind=k8) u(N/2+1),trigs(2*N),work(N)
    real(kind=k8) window,Dx,Nreal,mean,var
    complex(kind=k8) phi(N/2+1),phifd,im
    Nhalf = N/2
    NhalfOne = Nhalf + 1
    Nreal = N

    ! compute variance by the finite difference method

    call JUMPDIFFcharSym(0,h,phifd,lambda,sigmaZ,b,t)

    ! use symmetry property

    mean = aimag(phifd/h)
    var = abs(real(-(phifd - 2.0_k8 + conjg(phifd))/(h*h) - mean**2))
```

```
    ! fft window and frequences

    window = stdvs*sqrt(var)
    Dx = 2*window/Nreal

    call linSpace((pi/Dx),0.0_k8,NhalfOne,u)

    ! compute charteristic function

    call JUMPDIFFcharSym(N,u,phi,lambda,sigmaZ,b,t)

    ! inverse fft

    call zfftd(0,N,1.0,phi,dens,trigs,ifac,work,N,ierr)

    call zfftd(1,N,(1.0/(Nreal*Dx)),phi,dens,trigs,ifac,work,N,ierr)


    ! reorder datapoints
    xx(1:Nhalf) = dens(1:Nhalf)
    dens(1:Nhalf) = abs(dens(NhalfOne:N))
    dens(NHalfOne:N) = abs(xx(1:Nhalf))

    ! create abcissa

    call linSpace((-window),(window-Dx),N,xx)

end subroutine iFFTpdfinit

! computation of characteristic function

subroutine JUMPDIFFcharSym(N,u,phi,lambda,sigmaZ,b,t)
  integer, parameter:: k8 = selected_real_kind(15,150)
  real(kind=k8),parameter:: pi = 3.1415926535897932
  integer,intent(in):: N
  real(kind=k8),intent(in):: u(N/2+1)
  complex(kind=k8),intent(out):: phi(N/2+1)
  real(kind=k8),intent(in):: lambda
  real(kind=k8),intent(in):: sigmaZ
  real(kind=k8),intent(in):: b
  real(kind=k8),intent(in):: t
  real(kind=k8),parameter:: half = (1.0_k8)/(2.0_k8)
  complex(kind=k8),parameter:: im = (0.0_k8,1.0_k8)
  phi = exp(-t*half*(b*u)**2 + t*lambda*(1/(1-(-im*u*b/sigmaZ))-1))
end subroutine JUMPDIFFcharSym
```

B.2 A Split Step Backward Integration Code

This is the path integration code used in slightly altered forms for all the Langevin equations. Split step backward integration is used.

```
program LangevinPI
  !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
  ! Path integration for the stochastic differential equation
  !
  ! dXt = - a Xt dt + b dLt
  !
  ! written by Tore Selland Kleppe (toresell@stud.math.ntnu.no)
  ! feb 2006
  !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
   implicit none
   integer, parameter:: k8 = selected_real_kind(15,150)
   real(kind=k8), parameter:: pi = 3.1415926535897932_k8
   integer, parameter:: gridPoints = 101 ! number of point in the grid
   integer, parameter:: timeSteps = 500  ! number of timesteps
   integer, parameter:: Q = 2**7         ! number of FFTpoints
   integer, parameter:: intPoints = Q/2 ! number of integration points
   real(kind=k8), parameter:: T = (30.0_k8) ! max time
   real(kind=k8), parameter:: stdvs = 6.5_k8 ! st.dvs in integration
   real(kind=k8), parameter:: lambda = 0.0_k8 ! noise parameter
   real(kind=k8), parameter:: sigmaZ = 0.9_k8 ! noise  parameter
   real(kind=k8), parameter:: a = 0.5_k8 ! SDE parameter
   real(kind=k8), parameter:: b = 1.5_k8 ! SDE parameter
   real(kind=k8), parameter:: mu0 = 3.0_k8 ! initial density parameter
   real(kind=k8), parameter:: sigmaS0 = 0.2_k8 ! initial density parameter
   real(kind=k8) dt ! the timestep
   integer i,j,k  ! counter variables
   integer quarter
   real(kind=k8) abcis(gridPoints),tempPDF(gridPoints),time
   real(kind=k8) surf(gridPoints,(timeSteps+1))
   real(kind=k8) splineV(gridPoints),fftxx(Q),fftyy(Q)
   real(kind=k8) intGrid(intPoints),oldVals(intPoints),
   real(kind=k8) Euler(intPoints),Xstar(intPoints),TPD(intPoints)
   real(kind=k8) kern((intPoints)),test(intPoints+1)
   real(kind=k8) xmin,xmax,dx
   real(kind=k8) di,border,L
   real(kind=k8) mass,tempy,dummy1,dummy2
   real(kind=k8) mu0,sigmaS0
   real(kind=k8) temp,kk,mu,coeff

   dt = T/timeSteps
   write(*,*) 'timestep = ',dt
```

```
! set surface to zero

surf(1:gridPoints,1:timeSteps+1) = 0.0_k8

! initialize grid
xmin = -10.00_k8
xmax = 10.0_k8

call linSpaceDX(xmin,xmax,gridPoints,abcis,dx)

! initial distribution normal

call normalPDF(gridPoints,abcis,surf(:,1),mu0,sigmaS0)

! spline representation of initial density

call spline_cubic_set (gridPoints,abcis,surf(:,1),0,0.0_k8,1,0.0_k8,splineV)

! compute forward kernel used in the split step backward steps
! notice that we only use half the points

quarter = Q/4
write(*,*) gridPoints
call symJUMPDIFFpdfinit(Q,fftxx,fftyy,lambda,sigmaZ&
        ,b,dt,stdvs)

intGrid = fftxx((quarter+1):(3*quarter))
TPD = fftyy((quarter+1):(3*quarter))

! check inital distribution

mass = sum(surf(:,1))
mass = mass*dx
write(*,*) 'initial mass = ', mass

! master iteration loop
time = 0.0_k8
do i=2,timeSteps+1
   time = time + dt
   write(*,*) 'timestep # ',(i-1),' t = ',time

   ! grid Iteration loop
   do j=1,gridPoints

      ! integration iteration loop
      Xstar = intGrid + abcis(j)
      Euler = Xstar + dt*a*(Xstar)
      do k=2,(intPoints)
         if ((Euler(k) < xmax) .and. (Euler(k) > xmin)) then
```

```
                call spline_cubic_val(gridPoints,abcis,surf(:,(i-1)),splineV,&
                     Euler(k),oldVals(k),dummy1,dummy2)
             else
                oldVals(k) = 0.0_k8
             endif
          end do
          kern = oldVals*TPD
          dummy1 = sum(kern)*(Euler(2)-Euler(1))
          surf(j,i) = max(dummy1,0.0_k8)
       end do
       ! check /normalize mass
       mass = sum(surf(:,i))
       mass = mass*dx
       write(*,*) 'mass = ', mass
       surf(:,i) = ((1.0_k8)/mass)*surf(:,i)
       ! new spline surface
       call spline_cubic_set (gridPoints,abcis,surf(:,i),&
            0,0.0_k8,1,0.0_k8,splineV)
   end do
   call dumpResult(gridPoints,timeSteps,abcis,surf)

end program LangevinPI
```

```
!subroutine f1(n,X,Y,coeff)
!   integer, parameter:: k8 = selected_real_kind(15,150)
!   integer, intent(in):: n
!   real(kind=k8),intent(in):: X(n)
!   real(kind=k8),intent(out):: Y(n)
!   real(kind=k8),intent(in):: coeff
!   Y = coeff*X
!end subroutine f1
```

B.3 A Direct Integration Code

This is the path integration code used in slightly altered forms for all the Black-Scholes type equations. Direct integration is used.

```
program BSPI
  !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
  ! Path integration for the stochastic differential equation
  !
  ! dXt = a Xt dt + b Xt dLt
  !
  ! written by Tore Selland Kleppe (toresell@stud.math.ntnu.no)
  ! feb 2006
  !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
   implicit none
   integer, parameter:: k8 = selected_real_kind(15,150)
   real(kind=k8), parameter:: pi = 3.1415926535897932_k8
   integer, parameter:: gridPoints = 201 ! number of point in the grid
   integer, parameter:: timeSteps = 200  ! number of timesteps
   integer, parameter:: Q = 2**7         ! number of FFTpoints
   integer, parameter:: intPoints = Q/2 ! number of integration points
   real(kind=k8), parameter:: T = (1.0_k8) ! max time
   real(kind=k8), parameter:: stdvs = 40.5_k8 ! st.dvs in integration
   real(kind=k8), parameter:: a = 0.50_k8 ! SDE parameter
   real(kind=k8), parameter:: b = 0.1_k8 ! SDE parameter
   real(kind=k8), parameter:: X0 = 12.0_k8
   real(kind=k8) sigmaS0
   real(kind=k8) dt ! the timestep
   integer i,j,k  ! counter variables
   integer quarter
   real(kind=k8) abcis(gridPoints),tempPDF(gridPoints),time
   real(kind=k8) surf(gridPoints,(timeSteps+1))
   real(kind=k8) splineV(gridPoints),fftxx(Q),fftyy(Q)
   real(kind=k8) intGrid(gridPoints,intPoints),oldVals(intPoints)
   real(kind=k8) euro(gridPoints)
   real(kind=k8) TPD(gridPoints,intPoints),var(intPoints)
   real(kind=k8) kern(intPoints),test(intPoints+1)
   real(kind=k8) xmin,xmax,dx
   real(kind=k8) di,border,L
   real(kind=k8) mass,tempy,dummy1,dummy2
   real(kind=k8) mu0,sigmaS0
   real(kind=k8) temp,kk,mu,coeff,put,cal

   dt = T/timeSteps
   write(*,*) 'timestep = ',dt

   ! set surface to zero

   surf(1:gridPoints,1:timeSteps+1) = 0.0_k8

   ! initialize grid

   xmin = 0.01_k8
   xmax = 50.0_k8
```

```
call linSpaceDX(xmin,xmax,gridPoints,abcis,dx)

! initial distribution normal assuming deteministic initial cond.
! X0

sigmaS0 = b*b*dt
call normalPDF(gridPoints,abcis,surf(:,1),(1+a*dt)*X0,sigmaS0)

! initialize spline

call spline_cubic_set (gridPoints,abcis,surf(:,1),0,0.0_k8,1,0.0_k8,splineV)
! check inital distribution

mass = sum(surf(:,1))
mass = mass*dx
write(*,*) 'initial mass = ', mass

! compute backward transition densities
do j=1,gridPoints
   coeff = stdvs*sqrt(dt)*b*abcis(j)
   call linspace((abcis(j)-coeff),(abcis(j)+coeff)&
        ,intPoints,intGrid(j,1:intPoints))
   do k=1,intPoints
      if(intGrid(j,k)>0.0_k8)then
         TPD(j,k) = 1/(sqrt(2*pi*dt)*b*intGrid(j,k))
         TPD(j,k) = TPD(j,k)*exp(-(abcis(j)- (1+a*dt)*intGrid(j,k))**2&
              /(2*dt*(b*intGrid(j,k))**2))
      else
         TPD(j,k) = 0.0_k8
      endif
   end do
end do

! master iteration loop
time = 0.0_k8
do i=2,timeSteps+1
   time = time + dt
   write(*,*) 'timestep # ',(i-1),' t = ',time


   ! grid Iteration loop
   do j=1,gridPoints
      do k=2,(intPoints)
         if ((intGrid(j,k) < xmax) .and. (intGrid(j,k) > xmin)) then
            call spline_cubic_val(gridPoints,abcis,surf(:,(i-1)),splineV,&
                 intGrid(j,k),oldVals(k),dummy1,dummy2)
         else
            oldVals(k) = 0.0_k8
```

```
            endif
            kern(k) = oldVals(k)*TPD(j,k)
          end do

          ! trapeziodal integration rule

          kern(1:(intPoints-1)) = kern(1:(intPoints-1)) + kern(2:intPoints)
          dummy1 = 0.5_k8*sum(kern(1:(intPoints-1)))*(intGrid(j,2)-intGrid(j,1))
          surf(j,i) = max(dummy1,0.0_k8)
        end do

        ! check / normalize densities
        mass = sum(surf(:,i))
        mass = mass*dx
        write(*,*) 'mass = ', mass
        surf(:,i) = ((1.0_k8)/mass)*surf(:,i)

        ! new spline surface

        call spline_cubic_set (gridPoints,abcis,surf(:,i),&
             0,0.0_k8,1,0.0_k8,splineV)
    end do
    call dumpResult(gridPoints,timeSteps,abcis,surf)

end program BSPI
```