Diss. ETH No. 12566

Parameter Estimation and Bessel Processes in Financial Models and Numerical Analysis in Hamiltonian Dynamics

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Mathematics

presented by
ANJA GÖING-JAESCHKE
Dipl.-Math. Ruhr-Universität Bochum
born May 17, 1968
citizen of Bochum, Germany

accepted on the recommendation of
Prof. Dr. P. Embrechts, examiner
Prof. Dr. H. Hofer, co-examiner

1998

für Georg und meine Eltern

Acknowledgements

The work on this thesis was carried out at the Mathematics Department at ETH Zürich under the guidance of Prof. P. Embrechts and Prof. H. Hofer.

It is my pleasure to thank Prof. P. Embrechts for his invaluable advice and encouragement during the last years. He always managed to find time for helpful discussions. I am obliged to him for the opportunities to travel and for introducing me to other leading scientists in the field of applied probability and statistics.

I am very grateful to Prof. H. Hofer for giving me the opportunity to come to Zürich. His stimulating lectures, his interest in my work and lots of discussions were a continuous source of motivation and enthusiasm.

I take great pleasure in thanking Prof. M. Yor for introducing me to the theory of Bessel processes, guiding me through the literature and various technical difficulties and showing constant interest in the progress of my work on generalizations of Bessel processes.

My work on financial mathematics was financed by the "Forschungs-Interessengemeinschaft Risklab" and I take the opportunity to express my grateful thanks. The stimulating and fruitful discussions with members of Risklab deepened my knowledge of risk management and were a steady source of inspiration.

I also thank Prof. F. Delbaen for lots of interesting discussions. The material presented in Section 2.2 benefited from talks I had with Prof. M. Sørensen, Dr. B.M. Bibby and Dr. A.R. Pedersen and I much appreciated their help.

I would like to thank my friends and colleagues at the Mathematics Department for creating such a pleasant atmosphere.

Abstract

Multi-purpose parameter estimation methods play an increasingly important role in financial as well as insurance mathematics. In the first part of this thesis we study appropriate methods for estimating parameters in continuous time processes based on discrete observations, resulting in consistent and asymptotically normal estimators. Cox–Ingersoll–Ross (CIR) processes for modeling short-term interest rates are considered as a class of financial diffusion processes. Applying the estimation methods to simulated data of these processes we show the numerical quality of the resulting estimates.

Bessel processes emerge in many financial problems and have remarkable properties. Via a space-time transformation they are related to CIR processes. In many financial applications, the calculation of the first time a diffusion process reaches a certain level is important, as for instance in the case of barrier options. We explicitly calculate the Laplace transform of the first time Bessel processes and related processes, such as CIR processes, hit a given barrier. As a generalization, we introduce Bessel processes allowing negative dimensions. For instance, geometric Brownian motion with negative drift is closely related to these processes. We show their properties and derive their transition densities. Time reversal is a crucial tool throughout the calculations and time reversed diffusions are investigated in detail.

Turning to a different subject in the second part of this thesis, a numerical method is introduced for finding periodic orbits for Hamiltonian systems on a strictly convex energy surface. This strict convexity allows us to use a Legendre transform of the Hamiltonian in both variables simultaneously obtaining a minimizing problem on a constraint.

We introduce and describe this problem and prove the existence of a solution. The proposed numerical algorithm finds minima of this variational problem. A minimum found corresponds to a periodic orbit of minimal action of the original Hamiltonian system. The algorithm can not only be used to approximate periodic orbits of a Hamiltonian system but it also calculates symplectic invariants of open convex sets in phase space. These symplectic invariants for general symplectic manifolds, called symplectic capacities, were discovered by Ekeland, Hofer and Zehnder. We recall background material on symplectic invariants. The numerical program is applied to several examples of strictly convex bounded domains in $\mathbb{R}^{2n}$, assuming the boundary as the energy surface of a Hamiltonian function. Based on the results we discuss the properties of the numerical algorithm.

Kurzfassung

Vielseitig verwendbare Parameterschätzmethoden gewinnen immer mehr an Bedeutung sowohl in der Finanz- als auch in der Versicherungsmathematik. Im ersten Teil dieser Doktorarbeit untersuchen wir Methoden, die besonders dazu geeignet sind, Parameter in stetigen Prozesses basierend auf diskreten Beobachtungen zu schätzen. Die resultierenden Schätzer sind konsistent und asymptotisch normal. Als eine in der Finanzmathematik bedeutende Klasse von Diffusionsprozessen betrachten wir Cox–Ingersoll–Ross (CIR) Prozesse, mit denen Zinsraten modelliert werden. Durch Anwendung der Parameterschätzmethoden auf simulierte Daten dieser Prozesse wird die numerische Qualität der Schätzer gezeigt.

Besselprozesse tauchen in vielen Fragen in der Finanz auf und haben bemerkenswerte Eigenschaften. Durch eine Raum-Zeit-Transformation sind sie mit CIR Prozessen verbunden. In Finanzanwendungen ist oft der erste Zeitpunkt von Bedeutung, an dem Diffusionsprozesse ein vorgegebenes Niveau erreichen, zum Beispiel bei Barrier-Optionen. Wir berechnen explizit die Laplacetransformierte des ersten Zeitpunktes, an dem Besselprozesse und damit verbundene Prozesse wie CIR Prozesse auf eine gegebene Barriere treffen. Als eine Verallgemeinerung werden Besselprozesse eingeführt, die auch negativ-dimensional sein können. Geometrische Brownsche Bewegung mit negativem Drift ist zum Beispiel eng mit diesen Prozessen verbunden. Wir zeigen die Eigenschaften dieser Prozesse und leiten ihre Übergangsdichten her. In all den Berechnungen ist die Zeitumkehrung ein entscheidendes Werkzeug, und Diffusionsprozesse in umgekehrter Zeit werden im Detail untersucht.

Wir wenden uns im zweiten Teil dieser Doktorarbeit einem anderen

Thema zu. Eine numerische Methode wird eingeführt, die periodische Orbits von Hamiltonschen Systemen auf einer strikt konvexen Energiefläche bestimmt. Aufgrund der strikten Konvexität kann eine Legendretransformation der Hamiltonfunktion – in beiden Variablen gleichzeitig – benutzt werden. Damit erhalten wir ein Minimierungsproblem mit Nebenbedingungen. Wir leiten das Problem her und beweisen die Existenz einer Lösung. Der vorgestellte numerische Algorithmus findet Minima dieses Variationsproblems. Ein solches Minimum steht in Beziehung zu einem periodischen Orbit minimaler Aktion des ursprünglichen Hamiltonschen Systems. Der Algorithmus kann nicht nur dazu verwandt werden, periodische Orbits Hamiltonscher Systeme zu approximieren, sondern es können auch symplektische Invarianten von offenen konvexen Mengen des Phasenraumes berechnet werden. Diese symplektischen Invarianten für allgemeine symplektische Mannigfaltigkeiten, genannt symplektische Kapazitäten, wurden von Ekeland, Hofer und Zehnder entdeckt. Wir geben eine Einführung in symplektische Invarianten. Das numerische Programm wird auf mehrere Beispiele strikt konvexer beschränkter Gebiete im $\mathbb{R}^{2n}$ angewandt, wobei der Rand des Gebietes als Energiefläche der Hamiltonfunktion angenommen wird. Anhand der Ergebnisse werden die Eigenschaften des numerischen Algorithmus erörtert.

Contents

Acknowledgements **iii**

Abstract **v**

Kurzfassung **vii**

Parameter Estimation and Bessel Processes in Financial Models

1 Introduction **1**

2 Parameter Estimation **5**

2.1 Continuous Observations 6

2.2 Discrete Observations 12

2.2.1 Martingale Estimating Functions 14

2.2.2 Two Further Estimation Methods 23

3 The Cox–Ingersoll–Ross Model **29**

3.1 Boundary Behaviour 31

4 Parameter Estimation in CIR Processes 37

4.1 Estimation of Parameters in the Drift 38

4.2 Simulation Study in the Nonergodic Case 40

4.3 Estimation of Parameters in the Diffusion Coefficient . . 42

5 Some Generalizations of Bessel Processes 45

5.1 Main Properties of Bessel Processes 46

5.2 First Hitting Times of Radial Ornstein–Uhlenbeck Processes 50

5.3 BESQ Processes with Negative Dimensions and Extensions 62

5.4 Inhomogeneous Bessel Processes 76

5.5 Time Reversal 78

5.5.1 Doob's h-Transform 82

5.5.2 Checking a Time Reversal Theorem in Elworthy–Li–Yor 84

5.5.3 The Three-Parameters-Family of Processes with Law ${}^{\lambda}P^{\delta,\mu}$ 88

Bibliography 96

Parameter Estimation and Bessel Processes in Financial Models

Chapter 1

Introduction

Over the last few years various new derivative instruments have emerged in financial markets leading to a demand for versatile estimation methods for relevant model parameters. Typical examples include volatility, covariances and correlations. Moreover, questions of estimation concerning extremal events become more and more important in financial and insurance mathematics, for a recent detailed treatment of extreme value theory see Embrechts–Klüppelberg–Mikosch [27] and references therein.

Estimation of unknown parameters in a general diffusion process is discussed in Chapter 2. For continuously observed processes we develop the classical theory of maximum likelihood estimation including properties like consistency and asymptotic normality. In the case of discrete observations, the maximum likelihood estimator retains all the 'good' properties if the transition densities of the process are known in an explicit form. However, in most cases relevant for finance, we do not have explicit expressions for the underlying transition densities, and the use of approximate likelihood functions leads to inconsistent estimators when the time between observations is bounded away from zero. We discuss alternative estimation methods, especially martingale estimating functions proposed by Bibby–Sørensen [7], leading to consistent and asymptotically normal estimators.

The estimation methods discussed in Chapter 2 will be applied in Chapter 4 to a class of square root interest rate processes in order to investi-

gate the quality of the estimators. In Chapter 3 we study the behaviour of these square root diffusion processes. Cox, Ingersoll and Ross [16] proposed these processes to model short-term interest rates, known as Cox–Ingersoll–Ross (CIR) processes. In recent financial literature they are often studied from different points of view, or serve as important reference processes, see for instance McLeish–Kolkiewicz [69], Delbaen [20], Frydman [33], Deelstra–Parker [19], Duffie–Singleton [25], Chen–Scott [14] and Leblanc [66]. In Delbaen–Shirakawa [21] a more general square root diffusion model is investigated in detail, the constant elasticity of variance (CEV) model. This model allows for an additional exponential parameter in the diffusion coefficient and the behaviour of the process crucially depends on this parameter. We study the behaviour of a CIR process at its boundaries and show whether the process may hit a given barrier or not. This leads us to the question of when such processes will hit a given barrier for the first time. We deal with this problem in Chapter 5.

In Chapter 4 we apply martingale estimating functions and the discretized log-likelihood to simulated observations of CIR processes in order to explore the numerical performance of martingale estimating functions. In particular we investigate parameter estimation in diffusion coefficients. Since properties of the estimators are essentially only known for ergodic diffusions, we study the properties of the estimators for parameters in nonergodic CIR processes based on simulations.

Bessel processes play an important role for the CIR model for interest rates and we concentrate on this topic in Chapter 5. Bessel processes are a one-parameter family of diffusion processes that appear in many financial problems and have remarkable properties. As for their importance in financial mathematics, let us first mention their contribution to the problem of pricing Asian options with arithmetic asset average. A standard model for stock prices is geometric Brownian motion, i.e., the exponential of Brownian motion with drift. The feature that the exponential of Brownian motion with drift is a time-changed Bessel process is used by Geman–Yor [34] for inference in Asian option pricing. Another point emphasizing the relevance of Bessel processes for finance is that CIR processes are time-space-transformed squared Bessel processes, and can also be transformed by a time change to squared radial Ornstein–Uhlenbeck processes. Therefore, in order to obtain results for CIR processes, we will consider (squared) Bessel processes and (squared) radial Ornstein–Uhlenbeck processes, respectively.

In Chapter 5 we start by recalling the definition of Bessel processes and their basic properties. In financial mathematics, diffusion processes are often studied until they reach a certain level for the first time, as for instance in the case of barrier options (see e.g. Geman–Yor [35], Chesney–Geman–Jeanblanc-Picqué–Yor [15]). First hitting times are also very interesting for statistical reasons, e.g., if we want to condition a process on never hitting a given barrier. In Section 5.2 we deal with (the law of) the first time a squared radial Ornstein–Uhlenbeck process hits an arbitrary level – especially level zero. As the main result we give expressions of the Laplace transform of first hitting times by radial Ornstein–Uhlenbeck processes as well as by Bessel processes. In Section 5.3 we extend the class of Bessel processes to Bessel processes with possibly negative dimensions. These processes arise quite naturally when the exponential of Brownian motion with a negative drift is considered. We show their properties, derive their transition densities and, by a Girsanov transformation, extend the results to a wider class of processes.

As mentioned, in mathematical finance the exponential of Brownian motion with homogeneous drift is often used to model stock prices. In this context Bessel processes play an important role, since the exponential of Brownian motion with homogeneous drift is a time-changed Bessel process. As an extension we study the exponential of Brownian motion with inhomogeneous drift in Section 5.4. In the last section of Chapter 5 we investigate time reversal, which was an important tool in the previous chapters, in more depth. We review general results for time-reversed diffusions and as an application we check a time reversal result for radial Ornstein–Uhlenbeck processes.

Chapter 2

Parameter Estimation

In this chapter various methods for estimating parameters in a general diffusion process will be discussed. First we consider continuously observed diffusion processes and review the classical maximum likelihood theory. In the case of discrete observations we explain the limitation of the maximum likelihood method and discuss alternative estimation methods, as there are martingale estimating functions or a method based on approximating the transition densities. For a survey on estimation methods in discrete time models, as there are AR, ARCH, GARCH or HARCH models (see Engle [29], Bollerslev [9], Embrechts et al. [28]), as well as approximations of discrete time models via continuous time ones and vice versa we refer to [40].

We consider a probability space $(\Omega, \mathcal{A}, P)$ equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$ where $\mathcal{F}_0$ contains all sets of P-measure zero of $\mathcal{A}$. The general type of a diffusion process $X = (X_t)_{t\geq 0}$ considered is defined as the solution to the Itô stochastic differential equation

$$dX_t = b(t, X_t; \theta)\, dt + \sigma(t, X_t; \theta)\, dW_t, \quad X_0 = x_0, \quad t \geq 0, \tag{2.1}$$

where W is an r-dimensional $(\mathcal{F}_t)$-Wiener process, $\theta \in \Theta \subseteq \mathbb{R}^p$ a parameter, $b(\cdot, \cdot\,; \theta) : \mathbb{R}^+ \times \mathbb{R}^d \mapsto \mathbb{R}^d$ and $\sigma(\cdot, \cdot\,; \theta) : \mathbb{R}^+ \times \mathbb{R}^d \mapsto M^{d\times r}$, where $M^{d\times r}$ denotes the set of real $d \times r$ matrices. $(X_t)_{t\geq 0}$ is $(\mathcal{F}_t)$-adapted. In the following we always assume that the functions b and σ are Lipschitz continuous and that there exists a linear growth bound, assuring the

existence of a unique strong solution of (2.1), see Rogers–Williams [92], p. 128 and p. 136. X is a Markov process, see [92].

The situation will be discussed where a realization of the process X is observed, but the parameter θ is unknown to the observer. Hence, we have to construct sufficiently good estimators of θ and examine their properties. Deriving estimates of θ there are two different kinds of observations to distinguish. In Section 2.1 we consider continuous observations of X and recall the classical maximum likelihood theory which will be referred to at some points in Section 2.2 where we deal with parameter estimation methods for discretely observed diffusions.

2.1 Continuous Observations

Suppose that the diffusion process X satisfies

$$dX_t = b(\theta, X_t)dt + \sigma(\theta, X_t)dW_t, \quad X_0 = x_0, \quad t \geq 0, \tag{2.2}$$

where for convenience X and W are one-dimensional processes, b and σ are smooth functions, and the parameter $\theta \in \Theta \subseteq \mathbb{R}^p$ is to be estimated by continuous observations of X. P_θ denotes the law of the process X on the canonical space $C(\mathbb{R}_+, \mathbb{R})$ with the canonical filtration $\mathcal{F}_t = \sigma(X_s | s \leq t)$. Denote by $P_{\theta,t} := P_\theta | \mathcal{F}_t$ the restriction of P_θ to $\mathcal{F}_t$.

First we consider the case where $\sigma(\theta, x)$ does not depend on θ. Let $(\Omega, \mathcal{A}, \mu)$ be a probability space. A probability measure ν on $(\Omega, \mathcal{A})$ is absolutely continuous relative to μ if $\mu(A) = 0 \Rightarrow \nu(A) = 0$ for all $A \in \mathcal{A}$, and is denoted by $\nu << \mu$. The measures μ and ν are equivalent, denoted by $\mu \sim \nu$ if and only if $\nu << \mu$ and $\mu << \nu$. If σ does not vanish, the measures $P_{\theta,t}$ and $P_{\theta_1,t}$ for any $\theta, \theta_1 \in \Theta$, $t < \infty$, are equivalent, see Jacod [50] and Liptser–Shiryaev [68] §7. Then we are able to introduce the so called likelihood process

$$L_t^{\theta,\theta_1} = \frac{dP_{\theta,t}}{dP_{\theta_1,t}}, \tag{2.3}$$

the 'density process' of P_θ, relative to P_{θ_1}, or the 'Radon–Nikodym derivative'. The process $\left(L_t^{\theta,\theta_1}\right)$ is an $\mathcal{F}_t$-martingale (see Rogers–Williams [92], §4.17).

First we will concentrate on the case with a drift term linear in θ

$$dX_t = \theta b(X_t)dt + \sigma(X_t)dW_t,\ X_0 = x_0, \tag{2.4}$$

where b is possibly nonlinear, $\sigma > 0$, and the process (X_t) is observed in the time interval $[0, T]$. Then the likelihood process (2.3) in T with $\theta_1 = 0$ equals

$$L_T^{\theta,0} \equiv L_T(\theta) = \exp\left[\int_0^T \frac{\theta b(X_s)}{\sigma^2(X_s)} dX_s - \frac{1}{2}\int_0^T \frac{\theta^2 b^2(X_s)}{\sigma^2(X_s)} ds\right], \tag{2.5}$$

see e.g. Liptser–Shiryaev [68] §17, Kallianpur [52] p. 76, or Jacod [50]. In the following consider a constant diffusion term σ, say $\sigma \equiv 1$. For convenience we take the logarithm in (2.5) and obtain the logarithmic likelihood (log-likelihood) function

$$\ln L_T(\theta) = \int_0^T \theta b(X_s)dX_s - \frac{1}{2}\int_0^T \theta^2 b^2(X_s)ds. \tag{2.6}$$

Maximizing L_T, or equivalently $\ln L_T$, with respect to θ we obtain the so-called maximum likelihood estimator (MLE) $\hat{\theta}_T$ based on continuous observations of X in the interval $0 \le t \le T$. We remark that by definition of the likelihood function the measure P_{θ_1} with $P_\theta << P_{\theta_1}$ must not depend on an unknown parameter. Such a measure P_{θ_1} cannot be found, if besides the drift coefficient also the diffusion coefficient σ depends on an unknown parameter, i.e. this parameter cannot be estimated by the ML method.

The derivative with respect to θ of the log-likelihood is called the score function

$$\frac{\partial}{\partial\theta}\left(\ln L_T(\theta)\right) = \frac{L_T'(\theta)}{L_T(\theta)},$$

which in the case (2.6) has the form

$$\frac{\partial}{\partial\theta}\left(\ln L_T(\theta)\right) = \int_0^T b(X_s)dX_s - \theta\int_0^T b^2(X_s)ds.$$

Solving $\frac{\partial}{\partial\theta}(\ln L_T(\theta)) = 0$, we obtain the MLE

$$\hat{\theta}_T = \frac{\int_0^T b(X_s)dX_s}{\int_0^T b^2(X_s)ds}$$

and hence, using equality (2.4), we have

$$\hat{\theta}_T = \theta_0 + \frac{\int_0^T b(X_s)dW_s}{\int_0^T b^2(X_s)ds}, \tag{2.7}$$

where θ_0 denotes the true value of the parameter.

We derive some properties of the MLE $\hat{\theta}_T$. The *bias* in $\hat{\theta}_T$ as an estimator of θ_0 is defined as

$$b_{\hat{\theta}_T}(\theta_0) = \mathrm{E}_{\theta_0}(\hat{\theta}_T - \theta_0),$$

where E_{θ_0} denotes the expectation with respect to P_{θ_0}. Via (2.7) we have

$$E_{\theta_0}\hat{\theta}_T = \theta_0 + E_{\theta_0}\left[\frac{\int_0^T b(X_s)dW_s}{\int_0^T b^2(X_s)ds}\right]. \tag{2.8}$$

Note that in the case of a constant drift coefficient $b(X_t) \equiv b$

$$E_\theta \frac{bW_T}{b^2T} = \frac{1}{bT} E_\theta W_T = 0,$$

thus $E_{\theta_0}\hat{\theta}_T = \theta_0$, so that $\hat{\theta}_T$ is unbiased. In all other cases the estimator is biased, and in the case (2.6) we have under some regularity conditions the equality

$$b_{\hat{\theta}_T}(\theta_0) = \frac{d}{d\theta_0}\mathrm{E}_{\theta_0}\left(\int_0^T b^2(X_t)dt\right)^{-1},$$

see Liptser–Shiryaev [68], §17.2 and §17.3, Theorem 17.2 and 17.3.

Now we turn to the properties of consistency and asymptotic normality. Assuming some natural regularity conditions (see Ibragimov–Has'minskii [46] p. 83 and p. 90, Cramér [17] p. 500, or Kutoyants [62] p. 95), the stochastic differential equation (2.4) has an ergodic solution. For the ergodic theory we refer to e.g. Gihman–Skorokhod [38], §18. Then the stationary density $\tilde{p}$ of the process solves the (deterministic) stationary Fokker-Planck equation, which in the above case reduces to

$$\frac{d}{dx}\left[\theta b(x)\tilde{p}(x)\right] - \frac{1}{2}\frac{d^2}{dx^2}\tilde{p}(x) = 0.$$

Ergodicity implies

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T b(X_t)dW_t = 0 \quad P_{\theta_0}\text{ a.s.}$$

and

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T b^2(X_t)dt=\int_{-\infty}^{\infty}b^2(x)\tilde{p}(x)dx\quad P_{\theta_0}\text{ a.s.}\tag{2.9}$$

Hence for ergodic diffusions we conclude that it is possible to achieve ultimate arbitrary precision of the MLE $\hat{\theta}_T$ in (2.7) by infinitely increasing T, that means the MLE $\hat{\theta}_T$ is weakly consistent:

$$\lim_{T\to\infty}P_{\theta_0}\left[|\hat{\theta}_T-\theta_0|>\varepsilon\right]=0\tag{2.10}$$

for all $\varepsilon>0$. Furthermore, under these regularity conditions the MLE is asymptotically normal, that means as $T\to\infty$ the difference $\sqrt{T}(\hat{\theta}_T-\theta_0)$ is asymptotically normal with parameters $(0,\sigma^2(\theta_0))$, where

$$\sigma^2(\theta_0)=\left[\int_{-\infty}^{\infty}b^2(x)\tilde{p}(x)dx\right]^{-1},$$

see e.g. Kutoyants [62] p. 95, Kloeden–Platen [58] p. 243. We remark that this convergence is uniformly for all θ on a compact set $K\subseteq\Theta$, see e.g. Kutoyants [62] p. 94. Moreover, note that under the same regularity conditions the MLE is also consistent and asymptotic normal if the drift term is nonlinear (again see e.g. Kutoyants [62], p. 94). Moreover, we remark that $\sigma^2(\theta_0)$ equals $I^{-1}(\theta_0)$, where I denotes the so called Fisher information

$$\begin{aligned}I(\theta)&=\operatorname{Var}\left[\frac{\partial}{\partial\theta}\ln L_T(\theta)\right]\\&=\operatorname{E}\left[\left(\frac{\partial}{\partial\theta}\ln L_T(\theta)\right)^2\right],\end{aligned}$$

which can alternatively be computed by

$$I(\theta)=-\operatorname{E}\left[\frac{\partial^2}{\partial\theta^2}\ln L_T(\theta)\right],$$

see Lindgren [67] p. 249. Asymptotic normality can be used to determine confidence intervals for θ and to determine a 'suitable' value of T. Finally, among all possible estimators the MLE has the smallest mean squared error $\operatorname{E}[(\hat{\theta}-\theta_0)^2]=\operatorname{Var}\hat{\theta}+(\operatorname{E}\hat{\theta}-\theta_0)^2$, and the MLE is called efficient.

Example 1. a) The regularity conditions are satisfied in the case of observations coming from

$$dX_t = -\theta X_t dt + dW_t, \quad X_0 = 0,\ 0 \le t \le T, \tag{2.11}$$

where $\theta \in (\alpha, \beta)$, $\alpha > 0$, that is $-\theta < 0$ (see Kutoyants [62], p. 99). Thus the MLE $\hat{\theta}_T$ is consistent and asymptotically normal. The stationary density is

$$\tilde{p}(x) = \sqrt{\tfrac{\theta}{\pi}} e^{-\theta x^2},$$

and we have

$$\sigma^2(\theta_0) = \left[\sqrt{\tfrac{\theta_0}{\pi}} \int_{-\infty}^{\infty} x^2 e^{-\theta_0 x^2} dx \right]^{-1} = 2\theta_0,$$

and hence the MLE $\hat{\theta}_T$ is asymptotically normal with parameters $(0, 2\theta_0)$.

b) In the case $\theta \in (\alpha, \beta)$ and $\beta < 0$, that is $-\theta > 0$, the process has no stationary distribution. Nevertheless the MLE $\hat{\theta}_T$ can be shown to be consistent. Moreover, $e^{\theta_0 T}(\hat{\theta}_T - \theta_0)$ is asymptotically normal with parameters $(0, 4\theta_0^2)$ (see Kutoyants [62], p. 100).

c) In the case $\theta = 0$ the MLE $\hat{\theta}_T$ is consistent, but not asymptotically normal (see Kutoyants [62], p. 100).

We remark that in the discrete case an analogous distinction regarding parameter values is made.

After having treated the special case of a drift term linear in θ we give a short remark to the general case

$$dX_t = b(\theta, X_t)dt + dW_t,$$

where $b(\theta, x)$ may be a nonlinear function. By Taylor expansion for θ we have

$$b(\theta, x) = b(\theta_0, x) + (\theta - \theta_0)\frac{d}{d\theta} b(\theta_0, x) + O((\theta - \theta_0)^2),$$

where θ_0 again denotes the true value of the parameter θ. Hence $b(\theta, x)$ can be approximated by a term linear in θ. Since in practice a 'good guess' θ of θ_0 often is available, then with the above considerations it is

sufficient only to consider the already treated case: a drift term linear in θ.

In the following we give some remarks to the case where the diffusion term depends on an unknown parameter θ. As mentioned before in this case we cannot estimate θ by using maximum likelihood theory (see p. 7). As a first estimating problem we deal with the process

$$dX_t = \sigma(\theta)dW_t, \quad 0 \leq t \leq T. \tag{2.12}$$

In order to estimate θ we discretize the process and derive its limit. Consider partitions $\Delta_n = \left(t_0^{(n)}, \ldots, t_{m(n)}^{(n)}\right)$ of $[0,T]$ constructed in such a way that $\sum_{n=1}^{\infty} \sup_k |t_{k+1}^{(n)} - t_k^{(n)}| < \infty$ for $n \longrightarrow \infty$. Then for the resulting quadratic variation process we have the convergence result

$$\sum_{k=1}^{m} |W_{t_k^{(n)}} - W_{t_{k-1}^{(n)}}|^2 \longrightarrow T,$$

in probability for $n \longrightarrow \infty$, see Breiman [12] p. 262. Hence for (2.12) we obtain

$$\sum_{k=1}^{m} |X_{t_k^{(n)}} - X_{t_{k-1}^{(n)}}|^2 \longrightarrow T\sigma^2(\theta),$$

in probability for $n \longrightarrow \infty$, thus we are able to give an arbitrarily precise estimate of $\sigma^2(\theta)$. Notice that we do not obtain a direct estimate of θ.

If the unknown parameter θ splits into two parts (θ_1, θ_2) in the following way

$$dX_t = b(t, \theta_1)dt + \sigma(\theta_2)dW_t, \tag{2.13}$$

we are interested in estimating the part θ_2 in the diffusion coefficient. Denoting $\Delta X_k^{(n)} \equiv X_{t_k^{(n)}} - X_{t_{k-1}^{(n)}}$ and $\Delta W_k^{(n)} \equiv W_{t_k^{(n)}} - W_{t_{k-1}^{(n)}}$ we obtain by discretizing (2.13)

$$\begin{aligned}
&\sum_{k=1}^{m} \left[\left(\Delta X_k^{(n)}\right)^2 - \left(\sigma(\theta_2)\Delta W_k^{(n)}\right)^2 \right] \\
&= \sum_{k=1}^{m} b^2(t_k, \theta_1) \left(\Delta t_k^{(n)}\right)^2 + 2\sum_{k=1}^{m} b(t_k, \theta_1)\sigma(\theta_2)\Delta W_k^{(n)} \Delta t_k^{(n)},
\end{aligned}$$

which tends to 0 in probability as $n \longrightarrow \infty$. Thus the drift term is insignificant as $n \longrightarrow \infty$ and we are able to estimate $\sigma(\theta_2)$ just as in

the previous case. Then by treating θ_2 as known we can estimate θ_1 via the ML method. Finally, for the general problem with a diffusion coefficient also depending on X

$$dX_t = \sigma(\theta, X_t)dW_t, \tag{2.14}$$

a convergence in probability result is given by

$$\sum_{k=1}^{m} |X_{t_k^{(n)}} - X_{t_{k-1}^{(n)}}|^2 \longrightarrow \int_0^T \sigma^2(\theta, X_s)ds, \tag{2.15}$$

for $n \longrightarrow \infty$, see Feigin [30], Liptser–Shiryaev [68], §4. Again note that θ cannot be estimated directly. We remark that (2.15) can be used to obtain an estimate for the diffusion coefficient based on discrete observations $X_{t_0}, X_{t_1}, \ldots, X_{t_l}$ of the diffusion process X. But this estimation method usually leads to an extremely biased estimate of θ, unless the time between the observation points is very small, see e.g. Bibby [6]. In the next section parameter estimation based on discrete observations will be discussed in detail.

2.2 Discrete Observations

In this section we deal with the situation where the diffusion process X (2.1) has only been observed at not necessarily equally spaced discrete time points $0 = t_0 < t_1 < \ldots < t_n$. As for the estimation of parameters in discretely observed diffusions we refer to and partly follow Bibby and Sørensen [7, 8], Bibby [4, 3, 5, 6], Pedersen [84, 79, 83, 82, 80, 81], Kessler and Sørensen [57], Kessler [56, 55] and Kloeden, Schurz, Platen and Sørensen [59].

For the estimation of θ from discrete observations of X we have to distinguish between two cases. Either the transition densities of X are known explicitly or they are unknown, respectively known explicitly but numerically not easy to handle.

If the transition densities $p(s, x, t, y; \theta)$ of X are known, e.g. in the case of an Ornstein–Uhlenbeck process (see p. 18, Example 1), as an efficient estimator for θ we can choose the maximum likelihood estimator $\hat{\theta}_n$

which maximizes the likelihood function

$$L_n(\theta) = \prod_{i=1}^{n} p(t_{i-1}, X_{t_{i-1}}, t_i, X_{t_i}; \theta),$$

or equivalently the log-likelihood function

$$l_n(\theta) \equiv \ln L_n(\theta) = \sum_{i=1}^{n} \log\left(p(t_{i-1}, X_{t_{i-1}}, t_i, X_{t_i}; \theta)\right) \tag{2.16}$$

for θ, see e.g. Basawa [2], p. 14. In the case of time-equidistant observations ($t_i = i\Delta$, $i = 0, 1 \ldots, n$ for some fixed $\Delta > 0$) Dacunha-Castelle and Florens-Zmirou [18] prove consistency and asymptotic normality of $\hat{\theta}_n$ as $n \to \infty$, independent of the value of Δ. Unfortunately, in general the transition densities of X are unknown or numerically not easy to handle.

When the formulae of the transition densities of X cannot be used, the classical alternative estimator is obtained by approximating the log-likelihood function for θ based on continuous observations of X. For this log-likelihood function to be defined, the diffusion coefficient $\sigma(t, x; \theta) = \sigma(t, x)$ has to be known (see Section 2.1, p. 7). Under some regularity conditions (see Liptser–Shiryaev [68], §7) the log-likelihood function for θ based on continuous observations of X in $[0, t_n]$ can be written in terms of integrals

$$\begin{aligned} l^c_{t_n}(\theta) = &\int_0^{t_n} b(s, X_s; \theta)^T \left(\sigma(s, X_s)\sigma(s, X_s)^T\right)^{-1} dX_s \\ &-\frac{1}{2}\int_0^{t_n} b(s, X_s; \theta)^T \left(\sigma(s, X_s)\sigma(s, X_s)^T\right)^{-1} b(s, X_s; \theta) ds, \end{aligned} \tag{2.17}$$

and the usual approximation of these integrals leads to the approximate log-likelihood function for θ based on discrete observations of X

$$\begin{aligned} \tilde{l}_n(\theta) = &\sum_{i=1}^{n} b(t_{i-1}, X_{t_{i-1}}; \theta)^T \sigma_{i-1} \left(X_{t_i} - X_{t_{i-1}}\right) \\ &-\frac{1}{2}\sum_{i=1}^{n} b(t_{i-1}, X_{t_{i-1}}; \theta)^T \sigma_{i-1}\, b(t_{i-1}, X_{t_{i-1}}; \theta)(t_i - t_{i-1}), \end{aligned} \tag{2.18}$$

with the notation

$$\sigma_{i-1} \equiv \left(\sigma(t_{i-1}, X_{t_{i-1}})\, \sigma(t_{i-1}, X_{t_{i-1}})^T\right)^{-1}.$$

The question arises how parameters in the diffusion coefficient are to be estimated. If θ divides into two parts $\theta = (\theta_1, \theta_2)$, such that $b(\cdot, \cdot\,; \theta) = b(\cdot, \cdot\,; \theta_1)$ and $\sigma(\cdot, \cdot\,; \theta)$ is known up to the scalar factor θ_2, that means $\sigma(\cdot, \cdot\,; \theta) = \theta_2 \tilde{\sigma}(\cdot, \cdot)$, we may avoid the problem of parameter dependence. In this case θ_2^2 can be estimated in advance by a quadratic variance-like formula (see (2.15) and Florens-Zmirou [31]), and by inserting this estimate in $\sigma(\cdot, \cdot\,; \theta) = \theta_2 \tilde{\sigma}(\cdot, \cdot)$, the diffusion term can be assumed "known". Then the estimate of the approximate log-likelihood function $\tilde{l}_n$ can be used to estimate θ_1.

In the case where the diffusion coefficient $\sigma(\cdot, \cdot\,; \theta)$ depends on θ more generally, Hutton and Nelson [45] show that the discretized score function corresponding to $l_{t_n}^c$ can under certain regularity conditions still be used to estimate θ. That means in this case we are able to estimate θ based on discrete observations of X as well.

However, estimation methods for discrete observations that arise from the theory of continuous observations have the undesirable property that the estimators are strongly biased unless $\max_{1 \le i \le n} |t_i - t_{i-1}|$ is "small". If the time between observations is bounded away from zero, especially in the case of time-equidistant observations with Δ fixed, Florens-Zmirou [31] shows that the estimator $\tilde{\theta}_n$ of θ obtained by maximizing the approximate log-likelihood function $\tilde{l}_n(\theta)$ is inconsistent.

To overcome the difficulties regarding parameter dependence and the dependence of $\tilde{l}_n(\theta)$ on $\max_{1 \le i \le n} |t_i - t_{i-1}|$, we will propose different estimation approaches. First, in subsection 2.2.1, martingale estimating functions introduced by Bibby and Sørensen [7] are described. Martingale estimating functions are applied in Chapter 4. The basic ideas of the other two approaches in subsection 2.2.2 are to find good approximations to the transition densities, or respectively, to derive for a special class of diffusions the log-likelihood based on incomplete observations using the Kalman–Bucy filter.

2.2.1 Martingale Estimating Functions

The approach discussed below is proposed by Bibby and Sørensen [7, 8] and Bibby [3], see also Bibby [4, 5, 6]. We consider one-dimensional

diffusion processes defined by the stochastic differential equations

$$dX_t = b(X_t;\theta)\,dt + \sigma(X_t;\theta)\,dW_t, \tag{2.19}$$

where $X_0 = x_0$ and $t \geq 0$. Besides the usual assumptions on b and σ in (2.1), such that (2.19) has a unique solution for all θ in an open subset $\Theta \subseteq \mathbb{R}$, the functions b and σ are supposed to be twice continuously differentiable with respect to both arguments and σ is assumed to be positive. In contrast to (2.1), for convenience here we only consider the time-homogeneous case. To simplify the exposition further, assume that we can observe $\{X_t\}$ at discrete equidistant time points, say $\Delta, 2\Delta, \ldots, n\Delta$. Later we will give an extension to the case where X and θ are multi-dimensional.

Our goal is to estimate the parameter θ from these discrete observations $X_\Delta, X_{2\Delta}, \ldots, X_{n\Delta}$ of $\{X_t\}$. As mentioned in the introduction of Section 2.2, inference from discrete time observations is classically based on an approximation of the score function of the continuous log-likelihood function $l_t(\theta)$. Denote this approximation by $\dot{\tilde{l}}_n(\theta)$. For the definition of the continuous log-likelihood see Section 2.1, p. 7. In the one-dimensional and time-homogeneous case with σ not depending on θ, the continuous time log-likelihood function is

$$l_t(\theta) = \int_0^t \frac{b(X_s;\theta)}{\sigma^2(X_s)}\,dX_s - \frac{1}{2}\int_0^t \frac{b^2(X_s;\theta)}{\sigma^2(X_s)}\,ds, \tag{2.20}$$

see also (2.5) or (2.17). If we replace the Lebesgue integrals and the Itô integrals by Riemann-Itô sums and differentiate with respect to θ we get the approximate score function

$$\begin{aligned}\dot{\tilde{l}}_n(\theta) = &\sum_{i=1}^n \frac{\dot{b}(X_{(i-1)\Delta};\theta)}{\sigma^2(X_{(i-1)\Delta})}(X_{i\Delta} - X_{(i-1)\Delta}) \\ &-\Delta\sum_{i=1}^n \frac{b(X_{(i-1)\Delta};\theta)\,\dot{b}(X_{(i-1)\Delta};\theta)}{\sigma^2(X_{(i-1)\Delta})}\,.\end{aligned} \tag{2.21}$$

If σ depends on θ we use the same estimating function

$$\begin{aligned}\dot{\tilde{l}}_n(\theta) = &\sum_{i=1}^n \frac{\dot{b}(X_{(i-1)\Delta};\theta)}{\sigma^2(X_{(i-1)\Delta};\theta)}(X_{i\Delta} - X_{(i-1)\Delta}) \\ &-\Delta\sum_{i=1}^n \frac{b(X_{(i-1)\Delta};\theta)\,\dot{b}(X_{(i-1)\Delta};\theta)}{\sigma^2(X_{(i-1)\Delta};\theta)}\,.\end{aligned} \tag{2.22}$$

By using this approach for the estimation of θ the problem of inconsistency arises as already mentioned in the introduction of this section. To avoid this problem the basic idea is to modify the discretized score-function $\dot{\tilde{l}}_n(\theta)$ in such a way that a zero-mean P_θ-martingale is obtained. This martingale is used as an estimating function, a so-called martingale estimating function. Using martingale limit theorems, the estimator can be shown to be consistent and asymptotically normal, see Theorem 1.

We modify the discretized score function $\dot{\tilde{l}}_n$ and by subtracting its compensator, get a zero-mean P_θ-martingale with respect to the filtration defined by $\mathcal{F}_i = \sigma(X_\Delta, \ldots, X_{i\Delta})$, $i = 1, 2, \ldots$. The compensator is:

$$\sum_{i=1}^{n} E_\theta \left(\dot{\tilde{l}}_i(\theta) - \dot{\tilde{l}}_{i-1}(\theta) | \mathcal{F}_{i-1} \right)$$
$$= \sum_{i=1}^{n} \frac{\dot{b}(X_{(i-1)\Delta};\theta)}{\sigma^2(X_{(i-1)\Delta};\theta)} \left[(F(X_{(i-1)\Delta};\theta) - X_{(i-1)\Delta}) - \Delta\, b(X_{(i-1)\Delta};\theta) \right],$$

where

$$F(X_{(i-1)\Delta};\theta) \equiv E_\theta(X_{i\Delta} | X_{(i-1)\Delta}). \tag{2.23}$$

Thus we obtain a zero-mean martingale estimating function of the form

$$\tilde{G}_n(\theta) = \sum_{i=1}^{n} \frac{\dot{b}(X_{(i-1)\Delta};\theta)}{\sigma^2(X_{(i-1)\Delta};\theta)} \left(X_{i\Delta} - F(X_{(i-1)\Delta};\theta) \right). \tag{2.24}$$

More generally, we consider the class of zero-mean P_θ-martingale estimating functions

$$G_n(\theta) = \sum_{i=1}^{n} g_{i-1}(X_{(i-1)\Delta};\theta) \left(X_{i\Delta} - F(X_{(i-1)\Delta};\theta) \right), \tag{2.25}$$

where for $i = 1, \ldots, n$, the function g_{i-1} is $\mathcal{F}_{i-1}$-measurable and continuously differentiable in θ. We want to find the optimal estimating function within the class (2.25) in the asymptotic sense of Godambe and Heyde [39]. This function is within the class (2.25) in an L^2-sense closest to the score function $\dot{l}_n$ based on the usually unknown exact likelihood function. It is

$$G_n^*(\theta) = \sum_{i=1}^{n} \frac{\dot{F}(X_{(i-1)\Delta};\theta)}{\phi(X_{(i-1)\Delta};\theta)} \left(X_{i\Delta} - F(X_{(i-1)\Delta};\theta) \right), \tag{2.26}$$

where

$$\phi(X_{(i-1)\Delta};\theta) = E_\theta\left[(X_{i\Delta} - F(X_{(i-1)\Delta};\theta))^2 | X_{(i-1)\Delta}\right] \tag{2.27}$$

for $i = 1, \ldots, n$. The estimator obtained via the optimal estimating function G_n^* has a minimal asymptotic variance within the class (2.25).

It should be mentioned that for small Δ the martingale estimating function $\tilde{G}_n(\theta)$, defined in (2.24), is a first order approximation in Δ of G_n^*, that means $\tilde{G}_n(\theta)$ is approximately optimal.

In some cases there are possibly numerical problems in calculating $\dot{F}(x,\theta)$ in (2.26). One way to solve these problems involves approximating $\dot{F}(x,\theta)$ up to the order $O(\Delta^2)$. That leads to a martingale estimating function of the form

$$\begin{aligned} G_n^+(\theta) = \sum_{i=1}^n \Bigg[& \dot{b}(X_{(i-1)\Delta};\theta)\Delta + \frac{1}{2}\Delta^2 \Big(\dot{b}(X_{(i-1)\Delta};\theta) b'(X_{(i-1)\Delta};\theta) \\ & + b(X_{(i-1)\Delta};\theta)\dot{b}'(X_{(i-1)\Delta};\theta) + \frac{1}{2}(\dot{\sigma}^2(X_{(i-1)\Delta};\theta) b''(X_{(i-1)\Delta};\theta) \\ & + \sigma^2(X_{(i-1)\Delta};\theta)\dot{b}''(X_{(i-1)\Delta};\theta)) \Big) \Bigg] \frac{(X_{i\Delta} - F(X_{(i-1)\Delta};\theta))}{\phi(X_{(i-1)\Delta};\theta)} . \end{aligned} \tag{2.28}$$

As for the multi-dimensional case, suppose θ is k-dimensional, $\{X_t\}$ and $b(X_t;\theta)$ are d-dimensional, σ is a $d \times m$-dimensional matrix with $\sigma\sigma^T$ positive definite and the Wiener process $\{W_t\}$ is m-dimensional. Then the $k \times 1$-dimensional martingale estimating functions $\tilde{G}_n$ and G_n^* have the form

$$\begin{aligned} \tilde{G}_n(\theta) = \sum_{i=1}^n & \dot{b}(X_{(i-1)\Delta};\theta)^T \left(\sigma(X_{(i-1)\Delta};\theta)(\sigma(X_{(i-1)\Delta};\theta)^T\right)^{-1} \\ & \cdot \left(X_{i\Delta} - F(X_{(i-1)\Delta};\theta)\right) \end{aligned}$$

and

$$G_n^*(\theta) = \sum_{i=1}^n \dot{F}(X_{(i-1)\Delta};\theta)^T \phi(X_{(i-1)\Delta};\theta)^{-1} \left(X_{i\Delta} - F(X_{(i-1)\Delta};\theta)\right),$$

where ϕ is assumed to be positive definite and $\dot{b}$ and $\dot{F}$ denote the $d \times k$-dimensional matrices of partial derivatives with respect to the components of θ.

The asymptotic properties of the estimator $\hat{\theta}_n$ we obtain from the martingale estimating functions (2.24), (2.26) and (2.28), or more generally from the class of martingale estimating functions G_n of the form (2.25), are discussed by Bibby and Sørensen [7]. Under natural regularity conditions (see Bibby and Sørensen [7], pp. 7–9) we have for ergodic diffusions

Theorem 1 (Bibby–Sørensen [7]) An estimator $\hat{\theta}_n$, which solves the equation

$$G_n(\hat{\theta}_n) = 0,$$

exists with probability tending to one as $n \longrightarrow \infty$ under P_{θ_0}, where θ_0 denotes the true parameter value. Moreover, as $n \longrightarrow \infty$,

$$\hat{\theta}_n \longrightarrow \theta_0$$

in probability under P_{θ_0} and $\hat{\theta}_n$ is asymptotically normal in distribution under P_{θ_0}.

For the proof we refer to [7].

As a first example we consider the Ornstein–Uhlenbeck process where the transition densities are well-known.

Example 1
The Ornstein–Uhlenbeck process is the solution of the stochastic differential equation

$$dX_t = \theta X_t \, dt + \sigma \, dW_t, \tag{2.29}$$

with $X_0 = x_0$. In this case the drift coefficient is $b(x;\theta) = \theta x$ and the diffusion coefficient $\sigma(x,\theta) \equiv \sigma$ is assumed to be known. The transition probability is normal with mean $F(x;\theta) = x\, e^{\theta\Delta}$ and variance $\phi(\theta) = \frac{\sigma^2}{2\theta}(e^{2\theta\Delta} - 1)$. Hence the estimating function $\tilde{G}_n$ has the form

$$\tilde{G}_n(\theta) = \frac{1}{\sigma^2} \sum_{i=1}^{n} X_{(i-1)\Delta}(X_{i\Delta} - X_{(i-1)\Delta} e^{\theta\Delta}) \,,$$

and we obtain $\hat{\theta}_n$ as solution of $\tilde{G}_n(\theta) = 0$:

$$\hat{\theta}_n = \frac{1}{\Delta} \log \frac{\sum_{i=1}^{n} X_{(i-1)\Delta} X_{i\Delta}}{\sum_{i=1}^{n} X^2_{(i-1)\Delta}},$$

if $\sum_{i=1}^{n} X_{(i-1)\Delta} X_{i\Delta} > 0$. The estimators $\hat{\theta}_n$ we obtain from the martingale estimating functions G^+ and G^* are the same because here G^+ and G^* are proportional to $\tilde{G}$.

In the next example we consider a wider class of stochastic processes.

Example 2
The solutions of the stochastic differential equation

$$dX_t = (a + bX_t)\,dt + \psi(X_t)\,dW_t, \tag{2.30}$$

where $X_0 = x_0$ and the function ψ takes positive values in $\mathbb{R}$, are called mean-reverting processes. The unknown parameters are a and b. Cox–Ingersoll–Ross processes (3.2) discussed in Chapter 3, also known as square root diffusion processes, are special processes within this class. Our aim is to calculate the martingale estimating functions $\tilde{G}_n$ and G_n^*.

Lemma 1 *The function*

$$f(t) \equiv E_{a,b}(X_t|X_0)$$

solves

$$f'(t) = a + bf(t). \tag{2.31}$$

Proof: Write (2.30) in integral form

$$X_t = X_0 + \int_0^t (a + bX_s)ds + \int_0^t \psi(X_s)dW_s.$$

Conditioning on X_0 we have

$$\begin{aligned} \mathrm{E}_{a,b}(X_t|X_0) \;=\; & \mathrm{E}_{a,b}(X_0|X_0) + \mathrm{E}_{a,b}\left[\int_0^t (a + bX_s)ds|X_0\right] \\ & + \underbrace{\mathrm{E}_{a,b}\left[\int_0^t \psi(X_s)dW_s|X_0\right]}_{=0}, \end{aligned}$$

or equivalently

$$\mathrm{E}_{a,b}(X_t|X_0) = X_0 + at + b\int_0^t \mathrm{E}_{a,b}(X_s|X_0)ds,$$

and the claim follows. □

Corollary 1 *For*

$$F(X_{(i-1)\Delta}; a, b) = E_{a,b}\left(X_{i\Delta}|X_{(i-1)\Delta}\right)$$

we have

$$F(x; a, b) = xe^{b\Delta} + \frac{a}{b}(e^{b\Delta} - 1). \tag{2.32}$$

Proof: The solution of

$$f'(t) = a + bf(t), \quad f(t_0) = f_0,\ t \geq t_0,$$

is

$$f(t) = f_0\, e^{b(t-t_0)} + \frac{a}{b}(e^{b(t-t_0)} - 1).$$

Hence we have for $\mathrm{E}_{a,b}(X_{t_i}|X_{t_{i-1}}) \equiv f(t_i)$ with constant $\Delta \equiv t_i - t_{i-1}$ for all i

$$\begin{aligned}\mathrm{E}_{a,b}(X_{t_i}|X_{t_{i-1}}) &= \mathrm{E}(X_{t_{i-1}}|X_{t_{i-1}})\, e^{b\Delta} + \frac{a}{b}(e^{b\Delta} - 1) \\ &= X_{t_{i-1}} e^{b\Delta} + \frac{a}{b}(e^{b\Delta} - 1),\end{aligned}$$

and the claim follows. □

With (2.32) we are now able to calculate $\tilde{G}_n$ and G_n^*. Abbreviating

$$h_i(a, b) \equiv \left(X_{i\Delta} - X_{(i-1)\Delta}e^{b\Delta} + \frac{a}{b}(1 - e^{b\Delta})\right)$$

we have

$$\tilde{G}_n(a, b) = \left[\sum_{i=1}^{n} \frac{1}{\psi^2(X_{(i-1)\Delta})} h_i(a, b), \sum_{i=1}^{n} \frac{X_{(i-1)\Delta}}{\psi^2(X_{(i-1)\Delta})} h_i(a, b)\right]^T,$$

$$\begin{aligned}G_n^*(a, b) = \Bigg[&\sum_{i=1}^{n} \frac{e^{b\Delta} - 1}{b\phi(X_{(i-1)\Delta}; a, b)} h_i(a, b), \\ &\sum_{i=1}^{n} \frac{\Delta e^{b\Delta}(X_{(i-1)\Delta} + \frac{a}{b}) + \frac{a}{b^2}(1 - e^{b\Delta})}{\phi(X_{(i-1)\Delta}; a, b)} h_i(a, b)\Bigg]^T.\end{aligned}$$

Considering G^+ is not interesting since F is known.

The estimation equation $\tilde{G}_n(a,b) = 0$ can be solved explicitly. Abbreviating $\psi_{i-1}^2 \equiv \psi^2(X_{(i-1)\Delta})$ we obtain

$$e^{\tilde{b}_n\Delta} = \frac{\left(\sum_{i=1}^n \frac{X_{(i-1)\Delta}}{\psi_{i-1}^2}\right)\left(\sum_{i=1}^n \frac{X_{i\Delta}}{\psi_{i-1}^2}\right) - \left(\sum_{i=1}^n \frac{X_{(i-1)\Delta}X_{i\Delta}}{\psi_{i-1}^2}\right)\left(\sum_{i=1}^n \frac{1}{\psi_{i-1}^2}\right)}{\left(\sum_{i=1}^n \frac{X_{(i-1)\Delta}}{\psi_{i-1}^2}\right)^2 - \left(\sum_{i=1}^n \frac{X_{(i-1)\Delta}^2}{\psi_{i-1}^2}\right)\left(\sum_{i=1}^n \frac{1}{\psi_{i-1}^2}\right)}$$

and

$$\tilde{a}_n = \frac{\tilde{b}_n}{1 - e^{\tilde{b}_n\Delta}} \frac{\left(\sum_{i=1}^n \frac{X_{(i-1)\Delta}}{\psi_{i-1}^2}\right) e^{\tilde{b}_n\Delta} - \left(\sum_{i=1}^n \frac{X_{i\Delta}}{\psi_{i-1}^2}\right)}{\left(\sum_{i=1}^n \frac{1}{\psi_{i-1}^2}\right)}.$$

As regarding the martingale estimating function G_n^* we are able to find a closed expression for ϕ only in a few cases where the diffusion coefficient is rather simple. For instance, if the diffusion coefficient takes the form $\psi(x) = \sigma\sqrt{x}$ as in the Cox–Ingersoll–Ross model, see (3.2), we have with a similar argument as for F (that is, the conditional second moment solves an ordinary differential equation)

$$\phi(x;a,b) = \frac{\sigma^2}{2b^2}\left((a+2bx)e^{2b\Delta} - 2(a+bx)e^{b\Delta} + a\right).$$

We remark that in practical applications the estimating equations corresponding to G_n^* can be solved using a generalization of Newton's method, see Chapter 4.

Furthermore, if no explicit expressions for the conditional mean F and the conditional variance ϕ are known, then F and ϕ can be approximated by the sample mean and sample variance of a large number of simulated realizations of the diffusion process at the relevant time point.

So far, we considered estimation of parameters in the drift coefficient. If the unknown parameter θ is multi-dimensional and a part of θ is only found in the diffusion coefficient, then using the martingale estimating functions generated by the conditional mean, $\tilde{G}_n$ and G_n^*, leads to fewer estimation equations than parameters. That is, in this case neither G^* nor $\tilde{G}$ can be used to estimate the part of θ that only enters the diffusion coefficient. This disadvantage of $\tilde{G}_n$ and G_n^* motivates us to consider martingale estimating functions generated by higher order conditional

moments, for example by the conditional variance:

$$H_n(\theta)=\sum_{i=1}^{n} h(X_{(i-1)\Delta};\theta)\left[\left(X_{i\Delta}-F(X_{(i-1)\Delta};\theta)\right)^2-\phi(X_{(i-1)\Delta};\theta)\right], \tag{2.33}$$

with ϕ given by (2.27) and F given by (2.23). In analogy to (2.26) we obtain the optimal estimating function within the class (2.33), in the sense of Godambe and Heyde [39]. This optimal function takes the form

$$H_n^*(\theta)=\sum_{i=1}^{n} \frac{\dot{\phi}(X_{(i-1)\Delta};\theta)}{\psi(X_{(i-1)\Delta};\theta)}\left[\left(X_{i\Delta}-F(X_{(i-1)\Delta};\theta)\right)^2-\phi(X_{(i-1)\Delta};\theta)\right], \tag{2.34}$$

where ϕ is assumed to be differentiable in θ and ψ is the fourth conditional cumulant

$$\psi(X_{(i-1)\Delta};\theta)=E_\theta\left[\left(X_{i\Delta}-F(X_{(i-1)\Delta};\theta)\right)^4|X_{(i-1)\Delta}\right]-\phi(X_{(i-1)\Delta};\theta)^2,$$

where $i=1,\ldots,n$.

Combining the functions G_n and H_n leads to further martingale estimating functions. Consider the class

$$\begin{aligned} K_n(\theta) = \sum_{i=1}^{n} \big[& g(X_{(i-1)\Delta};\theta)\left(X_{i\Delta}-F(\theta)\right) \\ & +h(X_{(i-1)\Delta};\theta)\left(\left(X_{i\Delta}-F(\theta)\right)^2-\phi(\theta)\right)\big]. \end{aligned} \tag{2.35}$$

Following the optimal way of combining G_n and H_n described in Heyde [42], we obtain the optimal estimating function in the asymptotic sense of Godambe and Heyde [39] within the class (2.35) denoted by K_n^* with

$$\begin{aligned} g(\theta) &= \frac{\dot{\phi}(\theta)\eta(\theta)-F(\theta)\psi(\theta)}{\phi(\theta)\psi(\theta)-\eta^2(\theta)}, \\ h(\theta) &= \frac{\dot{F}(\theta)\eta(\theta)-\dot{\phi}(\theta)\phi(\theta)}{\phi(\theta)\psi(\theta)-\eta^2(\theta)}, \end{aligned} \tag{2.36}$$

where for abbreviation the first argument of all functions, that is $X_{(i-1)\Delta}$, has been left out and where η denotes the third conditional central moment

$$\eta(X_{(i-1)\Delta};\theta)=E_\theta\left[\left(X_{i\Delta}-F(X_{(i-1)\Delta};\theta)\right)^3|X_{(i-1)\Delta}\right],\ i=1,\ldots,n.$$

Often the functions F, ϕ, η and ψ have to be determined numerically by using simulations. A way around these numerical problems is to use the approximations

$$F(x,\theta) \doteq x + b(x,\theta)\,\Delta,$$
$$\phi(x,\theta) \doteq \sigma^2(x,\theta)\,\Delta,$$

where the error is of order $O(\Delta^2)$, see Florens-Zmirou [31]. With these approximations we obtain the approximately optimal function $\tilde{K}_n$ in the class (2.35) with

$$g(X_{(i-1)\Delta};\theta) = \frac{\dot{b}(X_{(i-1)\Delta};\theta)}{\sigma^2(X_{(i-1)\Delta};\theta)},$$
$$h(X_{(i-1)\Delta};\theta) = \frac{\dot{\sigma}\left(X_{i\Delta} - F(X_{(i-1)\Delta};\theta)\right)}{\sigma^3\left(X_{i\Delta} - F(X_{(i-1)\Delta};\theta)\right)\Delta}. \tag{2.37}$$

In the case of ergodic diffusions X, an analogous result to Theorem 1 for the class of martingale estimating functions G_n (2.25) can be shown for the class (2.35) of functions K_n based on the conditional mean and on the conditional variance, see Bibby [3, 7], assuring existence, consistency and asymptotic normality of the estimator $\hat{\theta}_n$ derived via K_n.

Summarizing, in contrast to the discretized log-likelihood function, with martingale estimating functions based on the conditional variance all parameters in the drift and in the diffusion coefficient as well can be estimated simultaneously. Together with the asymptotic properties as consistency and asymptotic normality independently of the time between observations, this shows that using martingale estimating functions is the appropriate method in the case of discrete observations.

In Chapter 4 we will estimate parameters in Cox–Ingersoll–Ross processes (3.2), see [16], also known as square root diffusion processes, via martingale estimating functions. We will compare the resulting estimates with the estimates obtained via the discretized log-likelihood (2.18).

2.2.2 Two Further Estimation Methods

Apart from the martingale estimating functions, other approaches are introduced in the literature, which lead under regularity conditions to

consistent and asymptotically normal estimators, see Pedersen [84, 83] and [79, 81]. The basic idea behind the first approach is to approximate the transition densities of the diffusion process X in (2.1). For a special class of diffusions the log-likelihood based on incomplete observations is derived in the second approach, using the Kalman–Bucy filter. For a detailed summary of these two approaches see [40].

Approximation of the transition densities of X by a sequence of transition densities of approximating Markov processes

The crucial point of the approach introduced by Pedersen in [84, 83] is to approximate the transition densities $p(s,x,t,y;\theta)$ of X by a sequence of transition densities $(p_N(s,x,t,y;\theta))_{N=1}^{\infty}$ of approximating Markov processes and then to define the approximate log-likelihood functions

$$l_{n,N}(\theta) \equiv \sum_{i=1}^{n} \log\left(p_N(t_{i-1}, X_{t_{i-1}}, t_i, X_{t_i}; \theta)\right). \tag{2.38}$$

The approximating Markov processes are derived by using the Euler approximation. For fixed $0 \le s < t$, $x \in \mathbb{R}^d$, $\theta \in \Theta$ and $N \in \mathbb{N}$, we consider for $k = 0, 1, \ldots, N$:

$$\begin{aligned} \tau_k &= s + k\,\frac{t-s}{N} \\ Y_s^{(N)} &= x \\ Y_{\tau_k}^{(N)} &= Y_{\tau_{k-1}}^{(N)} + \frac{t-s}{N}\, b\left(\tau_{k-1}, Y_{\tau_{k-1}}^{(N)}; \theta\right) \\ &\qquad + \sigma\left(\tau_{k-1}, Y_{\tau_{k-1}}^{(N)}; \theta\right)\left(W_{\tau_k}^{\theta,s} - W_{\tau_{k-1}}^{\theta,s}\right). \end{aligned}$$

Under very strict regularity conditions it can be shown, see Pedersen [84], that the transition densities $(p_N(s,x,t,y;\theta))_{N=1}^{\infty}$ of these Markov processes obtained via Euler approximation converge to the transition densities $p(s,x,t,y;\theta)$ of X as N tends to infinity.

The sequence $(l_{n,N}(\theta))_{N=1}^{\infty}$ builds a connection between $\tilde{l}_n(\theta)$ (see (2.18)) and $l_n(\theta)$ in the following sense: the approximation $l_{n,1}(\theta)$ is a generalization of $\tilde{l}_n(\theta)$ with no restrictions on $\sigma(\cdot,\cdot\,;\theta)$ regarding parameter dependence, each $l_{n,N}(\theta)$ for $N \ge 2$ is an improvement of $l_{n,1}(\theta)$, and as N tends to infinity, for each $\theta \in \Theta$ and $n \in \mathbb{N}$, $l_{n,N}(\theta)$ converges

in probability under P_{θ_0} to $l_n(\theta)$, where θ_0 denotes the true parameter value.

As shown in Pedersen [83], the maximum likelihood estimator $\hat{\theta}_{n,N}$ obtained by maximizing $l_{n,N}(\theta)$ is consistent and asymptotically normal as n and N tend to infinity.

After these theoretical considerations we deal with the actual calculation of $l_{n,N}(\theta)$ for large values of N. In order to maximize $l_{n,N}(\theta)$, numerical algorithms usually require the value of $l_{n,N}(\theta)$ in a finite number of points θ. Since p_1 has a closed expression, see Pedersen [84], $l_{n,N}(\theta)$ is explicitly given for $N = 1$. But for $N \geq 2$ this is in general not the case and the densities p_N are approximated by

$$\frac{1}{M} \sum_{m=1}^{M} p_1 \left(\tau_{N-1}, Y_{N-1}^m, t, y; \theta \right) , \tag{2.39}$$

see Pedersen [84], for M chosen sufficiently large and where Y_{N-1}^m is the m th via the Euler scheme simulated value of $Y_{\tau_{N-1}}^{(N)}$, $m = 1, \ldots, M$.

Note that in order to calculate $p_N(s, x, t, y; \theta)$ these simulations have to be done for all needed values in (2.38) of $0 \leq s < t$, $x, y \in \mathbb{R}^d$ and $\theta \in \Theta$, for $k = 0, 1, \ldots, N-1$ and $m = 1, \ldots, M$. Therefore, the numerical maximization of $l_{n,N}(\theta)$ means an extremely high computational effort. For an extensive simulation study of this method see Honoré [44]. Apart from the fact that theoretical results about the estimators are only proved under very strict regularity conditions, this is the convincing reason for giving preference to the martingale estimating functions approach in Section 2.2.1.

Deriving the log-likelihood based on incomplete observations by means of the Kalman–Bucy filter

In the approach proposed by Pedersen [79, 81] diffusion processes defined by

$$dX_t = (A_t X_t + a_t)\, dt + B_t\, dW_t,\ X_0 = x_0,\ t \geq 0, \tag{2.40}$$

are considered, where $A : [0, \infty) \mapsto M^{d \times d}$, $a : [0, \infty) \mapsto \mathbb{R}^d$ and $B : [0, \infty) \mapsto M^{d \times m}$ $(d \leq m)$ are deterministic functions of t, known up to a parameter $\theta \in \Theta \subseteq \mathbb{R}^p$, and W is an m-dimensional Wiener process. Stochastic differential equations of the form (2.40) are known as linear

in the narrow sense. Their solutions are Gaussian diffusion processes, i.e. all joint distributions are Gaussian, and hence, the explicit form of the transition densities is known. Thus, if the process X is completely observable, i.e. if we can observe all coordinates of X at time points $0 = t_0 < t_1 < \ldots < t_n$, we can use as usual the log-likelihood function to obtain the maximum likelihood estimator for the parameter θ, see (2.16), and there is no need to search for a different method. The approach by Pedersen [79, 81] is useful, if estimates of θ have to be derived from possibly incomplete discrete observations of X in (2.40) at time points $0 = t_0 < t_1 < \ldots < t_n$.

Before giving a short summary of the approach [79, 81] we make a few remarks. In Chapter 3 and 4 we do not deal with the problem of incomplete observations, and being restricted to a class of diffusions defined by (2.40) we cannot apply this method to the processes discussed in Chapter 3. However, since the idea of deriving the log-likelihood based on possibly incomplete observations using the Kalman–Bucy filter is very interesting, we very briefly sketch this approach.

It is shown, see Pedersen [79], that the Markov chain $(X_{t_i})_{i=0}^n$ is a special case of the stochastic system

$$X_i = D_i X_{i-1} + S_i + \varepsilon_i,\ i = 1, \ldots, n, \tag{2.41}$$

where $(X_i)_{i=0}^n$ are random $d \times 1$ vectors, $(D_i)_{i=0}^n$ are non-random $d \times d$ matrices, $(S_i)_{i=1}^n$ are non-random $d \times 1$ vectors, $X_0 \sim N_d(x_0, V_0)$, $\varepsilon_i \sim N_d(0, V_i)$, $i = 1, \ldots, n$ and $X_0, \varepsilon_1, \ldots, \varepsilon_n$ are stochastically independent. The non-random elements in (2.41) including x_0 and $(V_i)_{i=0}^n$ are given up to the parameter $\theta \in \Theta \subseteq \mathbb{R}^p$. We assume the observable quantities are $Y_0, Y_1, \ldots, Y_n$ given by

$$Y_i = T_i X_i + U_i + e_i,\ i = 0, 1, \ldots, n, \tag{2.42}$$

where $(T_i)_{i=0}^n$ are non-random $k \times d$ matrices ($k \leq d$), $(U_i)_{i=0}^n$ are non-random $k \times 1$ vectors, $e_i \sim N_k(0, W_i)$ and $X_0, \varepsilon_1, \ldots, \varepsilon_m, e_0, e_1, \ldots, e_n$ are stochastically independent, $i = 0, 1, \ldots, n$. The matrices (T_i) specify the observable parts of (X_i), the vectors (U_i) are other inputs and the vectors (e_i) are measurement errors. It is obvious that both the case of complete observations, $Y_i = X_i$, and the case of partial observations without measurement errors are contained in the general case (2.42).

In the case where all non-random elements in (2.41) and (2.42) including x_0, $(V_i)_{i=0}^n$ and $(W_i)_{i=0}^n$ are known, the conditional expectations

$(\mathrm{E}(X_i|Y^i))_{i=0}^n$, respectively $(\mathrm{E}(X_i|Y^{i-1}))_{i=1}^n$, are the best predictors of X_i given the observations $Y^i \equiv (Y_0^T, \ldots, Y_i^T)$, respectively given $Y^{i-1} \equiv (Y_0^T, \ldots, Y_{i-1}^T)^T$, in the sense of minimal variance. They can be calculated by means of an iterative procedure referred to as the Kalman–Bucy filter. The Kalman–Bucy filter and filtering theory in general are well-studied for continuous time stochastic processes with continuous observations (see Liptser and Shiryaev [68], Kallianpur [52] and Øksendal [77]). Regarding the Kalman–Bucy filter for discrete time stochastic processes, we refer to Brockwell–Davis [13]. The important point is that besides the predictors $(\mathrm{E}(X_i|Y^i))$ and $(\mathrm{E}(X_i|Y^{i-1}))$ the Kalman–Bucy filter gives the density $p_{i|i-1}$ for the conditional distribution of Y_i given Y^{i-1}, i.e. we are able to calculate the transition densities of Y.

If the non-random elements in (2.41) and (2.42) including x_0, $(V_i)_{i=0}^n$ and $(W_i)_{i=0}^n$ are known up to a parameter $\theta \in \Theta \subseteq \mathbb{R}^p$, this means, that the log-likelihood function for θ based on observations of $Y_0, Y_1, \ldots, Y_n$ can be calculated, from which an estimation of θ can be obtained. A detailed numerical discussion can be found in Pedersen [81] in the case where the deterministic functions A, a and B in (2.40) are not completely known.

Chapter 3

The Cox–Ingersoll–Ross Model

In the previous chapter parameter estimation methods in continuous time models based on discrete observations as martingale estimating functions were studied in detail. In the next chapter these estimation methods will be applied to the class of square root diffusion models also known as Cox–Ingersoll–Ross models, see (3.2). As emphasized in the Introduction, this class of models is an often studied relevant one in mathematical finance literature.

The stochastic differential equation

$$dX_t = (a + b\,X_t)\,dt + \sigma\,\sqrt{|X_t|}\,dW_t\,,\ X_0 = x_0\,, \tag{3.1}$$

with a, $b \in \mathbb{R}$, $\sigma > 0$ and (W_t) standard Brownian motion, has on $[0, \infty)$ for every given initial value $X_0 = x_0$ a unique strong solution (X_t) (Revuz–Yor [89] § IX.3, see also Ikeda–Watanabe [47], p. 221). If $a = 0$ and $x_0 = 0$, the solution of (3.1) is $X_t \equiv 0$. From the comparison theorem (see Revuz–Yor [89] Theorem IX.(3.7)) we deduce $X_t \geq 0$ for the case $a \geq 0$, $x_0 \geq 0$; hence, in this case the absolute value in (3.1) can be omitted. Whether the process (3.1) hits 0 will be discussed in Section 3.1 where the boundary behaviour of the process is described in detail. Let us further remark that in the case $a < 0$ these processes correspond to squared Bessel processes with so-called negative dimensions which

are introduced in Section 5.3. As for the transformation of solutions of (3.1) to squared Bessel processes see formula (5.8); note that (5.8) is also valid for $a < 0$.

Cox–Ingersoll–Ross [16] introduced this square root diffusion model

$$dX_t = (a + b\,X_t)\,dt + \sigma\,\sqrt{X_t}\,dW_t\,,\ X_0 = x_0\,, \tag{3.2}$$

with parameters $a \geq 0$, $b \in \mathbb{R}$ and $\sigma > 0$ as a model for interest rates. In this section we discuss the properties of the solution of (3.2), called the *Cox–Ingersoll–Ross* (CIR) *process*.

If the process will hit 0 almost surely the question arises *when* the process will hit 0 for the first time, i.e. we are interested in the distribution of

$$T_0 = \inf\{t | X_t = 0\},$$

the first hitting time of 0. We discuss first hitting times by CIR processes, or more precisely by squared radial Ornstein–Uhlenbeck processes, which are essentially CIR processes (see Section 5.1, I.6), in Section 5.2.

The transition density of the CIR process (3.2) is

$$p_t(x,y) = \frac{2b\,e^{-bt}}{\sigma^2(1-e^{-bt})}\left(\frac{y}{xe^{bt}}\right)^{\frac{\nu}{2}} \exp\left[-\frac{2b}{\sigma^2}\,\frac{x+e^{-bt}y}{1-e^{-bt}}\right] I_\nu\left[\frac{4b\sqrt{xy}\,e^{-\frac{b}{2}t}}{\sigma^2(1-e^{-bt})}\right],$$

where $\nu = \frac{2a}{\sigma^2} - 1$, and I_ν is the modified Bessel function of the first kind of order ν

$$I_\nu(x) = \sum_{n=0}^{\infty} \frac{1}{n!\,\Gamma(\nu+n+1)}\left(\frac{x}{2}\right)^{2n+\nu}.$$

Consider the case $a > 0$. Following Karlin–Taylor [53], p. 220, there exists a unique stationary density only if $b < 0$

$$p(x) = \left(\frac{-2b}{\sigma^2}\right)^{\frac{2a}{\sigma^2}} \frac{1}{\Gamma\left(\frac{2a}{\sigma^2}\right)}\, x^{\frac{2a}{\sigma^2}-1}\, e^{\frac{2bx}{\sigma^2}}.$$

In this case the CIR process has a *mean reverting tendency*, that is, it tends away from extremely high or low values and reverts to the long-term mean $-\frac{a}{b}$. As an illustration see the simulated paths in Figures 3.1 and 3.2.

Stationary CIR processes have also the ergodic property, see Kunita [61], which means that time average equals ensemble average, see (2.9) and e.g. Wong–Hajek [100] §2.6.

3.1 Boundary Behaviour

In this section we are interested in the behaviour of the CIR process (3.2) at its left boundary $l = 0$ and at its right boundary $r = \infty$. We follow the classifications of boundary behaviour developed in Karlin–Taylor [53] § 15.6-8.

We use the terminology of Feller: A boundary is called *regular*, if the diffusion process can both enter and leave the boundary. From the interior of the state space an *entrance* boundary can never be reached by the diffusion process but it can be started there. A boundary is *natural*, if the diffusion process can neither reach it in finite time nor can be started from it. The diffusion process can reach an *exit* boundary, but starting at an exit boundary it is impossible to reach any interior state. Consider the left boundary l; the notations for the right boundary are analogous. An *attracting* left boundary l can be reached prior to reaching an arbitrary state b with positive probability from any interior starting point $x < b$, but not necessarily in finite time. Such a boundary does not have to be in the state space of the diffusion process. An attracting left boundary l is *attainable* if the first hitting time of l by the diffusion process starting in $x > l$ is finite with positive probability.

In order to classify the boundary behaviour of a diffusion process, functionals that measure e.g. the speed of the process near the boundary or the time it takes to reach an interior point starting at the boundary, have to be calculated; for details see Karlin–Taylor [53] § 15.6.

The behaviour of the CIR process with $a > 0$ at its left boundary $l = 0$ is summarized in Table 3.1. For $a < \frac{\sigma^2}{2}$ the boundary $l = 0$ is (instantaneously) reflecting, that is, when a sample path reaches 0, then it returns immediately to the interior of the state space in a reflecting manner; the Lebesgue measure of the set $\{t|X_t = 0\}$ is zero. For all $b \in \mathbb{R}$ the boundary $r = \infty$ is *natural* and *unattainable*. Furthermore,

Criteria	Behaviour
$a < \frac{\sigma^2}{2}$	regular, reflecting, attracting, attainable $b \leq 0$: $P(\inf\{t > 0 \mid X_t = 0\} < \infty) = 1$ $b > 0$: $P(\inf\{t > 0 \mid X_t = 0\} < \infty) \in]0, 1[$
$a \geq \frac{\sigma^2}{2}$	entrance boundary, nonattracting, unattainable

Table 3.1: *Behaviour of the CIR process ($a > 0, \sigma > 0$) at the boundary $l = 0$*

we have

$$r \text{ is } \begin{cases} \textit{attracting}, \text{ if } b > 0 \text{ or } (b = 0 \text{ and } a > \frac{\sigma^2}{2}) \\ \textit{nonattracting}, \text{ if } b < 0 \text{ or } (b = 0 \text{ and } a \leq \frac{\sigma^2}{2}) \,. \end{cases} \tag{3.3}$$

In the case $a = 0$, note that as soon as the CIR process reaches 0 it stays 0, i.e. 0 is an exit boundary. More precisely, for $a = 0$ the process will become *extinct*, also called *absorbed*, almost surely for $b \leq 0$, and with positive probability for $b > 0$. From the behaviour of the process with $a = 0$ we deduce via the comparison theorem (Revuz–Yor [89], Theorem IX.(3.7)) that the processes with $a < 0$ hit 0 a.s. in the case $b \leq 0$.

Figures 3.1, 3.2, 3.3 and 3.4 show simulated paths of the CIR process in four cases, which illustrate the different boundary behaviour of CIR processes depending on the parameters. The simulations were made by using the 1.5 strong Taylor scheme. This scheme is obtained by adding more terms from the Itô–Taylor expansion to the Milstein scheme, lead-

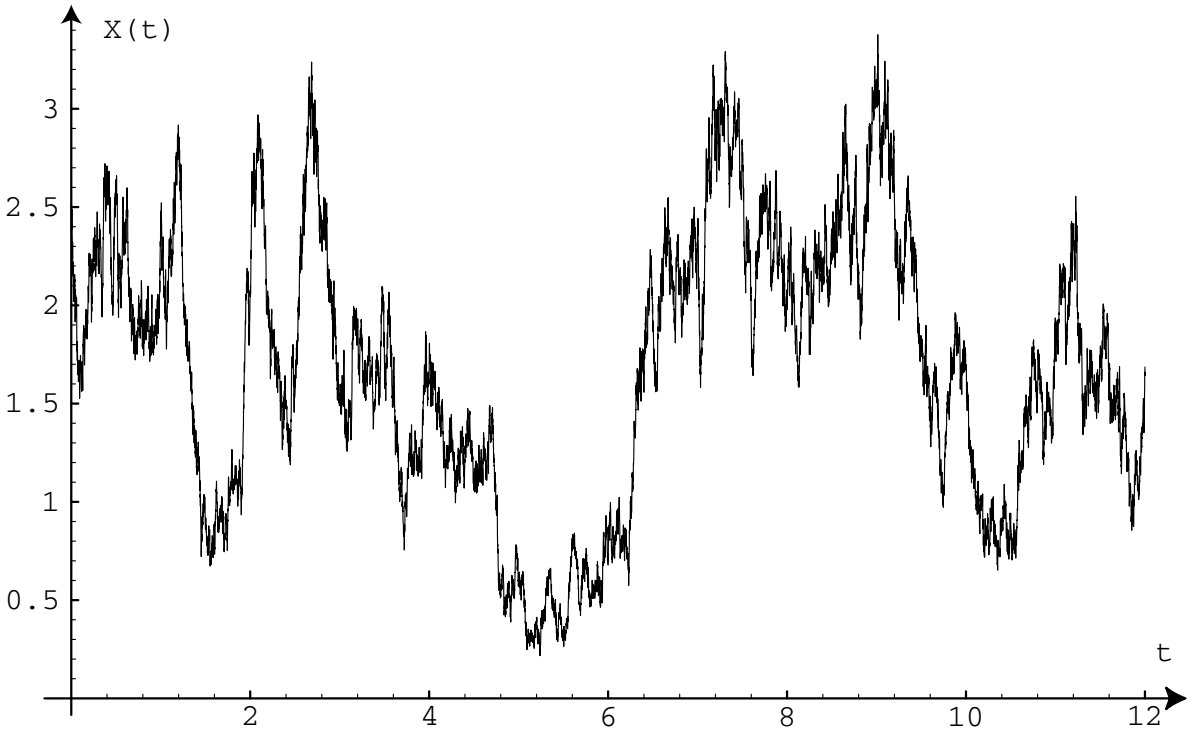

Figure 3.1: *A simulated path of the CIR process with* $x_0 = 2.5$, $a = 1$, $b = -0.5$, $\sigma = 1$.

ing to a strong order of convergence 1.5:

$$\begin{aligned} X_{(n+1)\Delta} = X_{n\Delta} + \Delta\,(a + b\,X_{n\Delta}) + \sigma\,\sqrt{X_{n\Delta}}\,\Delta W + \frac{\sigma^2}{4}\,((\Delta W)^2 - \Delta) \\ + b\,\sigma\,\sqrt{X_{n\Delta}}\,\Delta Z + \frac{1}{2}\,\Delta^2 b\,(a + b\,X_{n\Delta}) \\ + \frac{\sigma}{2\sqrt{X_{n\Delta}}}\left(a + bX_{n\Delta} - \frac{\sigma^2}{4}\right)(\Delta W\,\Delta - \Delta Z), \end{aligned} \tag{3.4}$$

where $\Delta W = \sqrt{\Delta}U_1$, $\Delta Z = \Delta^{3/2}(U_1 + U_2/\sqrt{3})/2$, and U_1 and U_2 are independent $N(0,1)$-distributed random variables. For more details see Kloeden–Platen [58] § 10.4.

309.83342pt

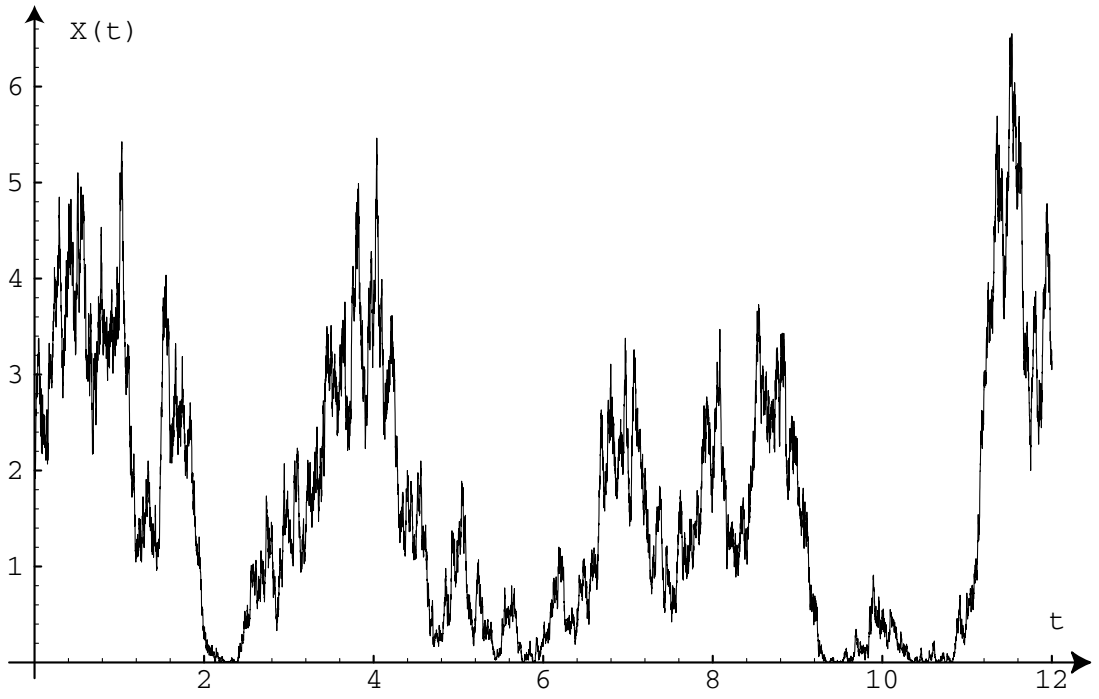

Figure 3.2: *A simulated path of the CIR process with* $x_0 = 2.5$, $a = 1$, $b = -0.5$, $\sigma = 2$.

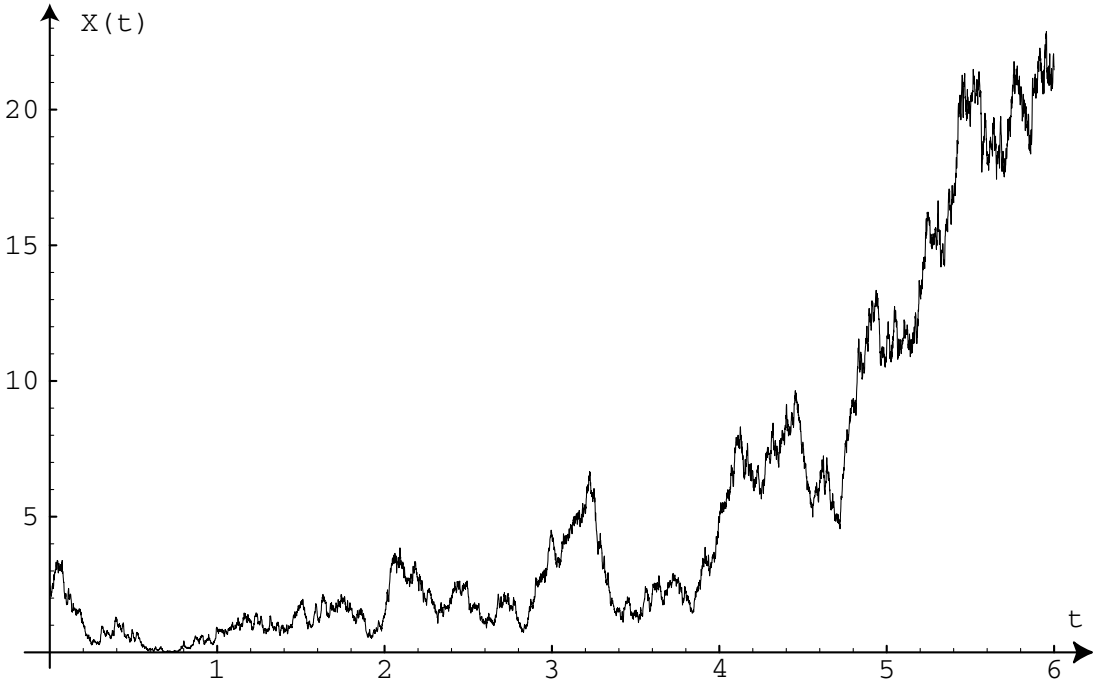

Figure 3.3: *A simulated path of the CIR process with* $x_0 = 2.5$, $a = 1$, $b = 0.5$, $\sigma = 2$.

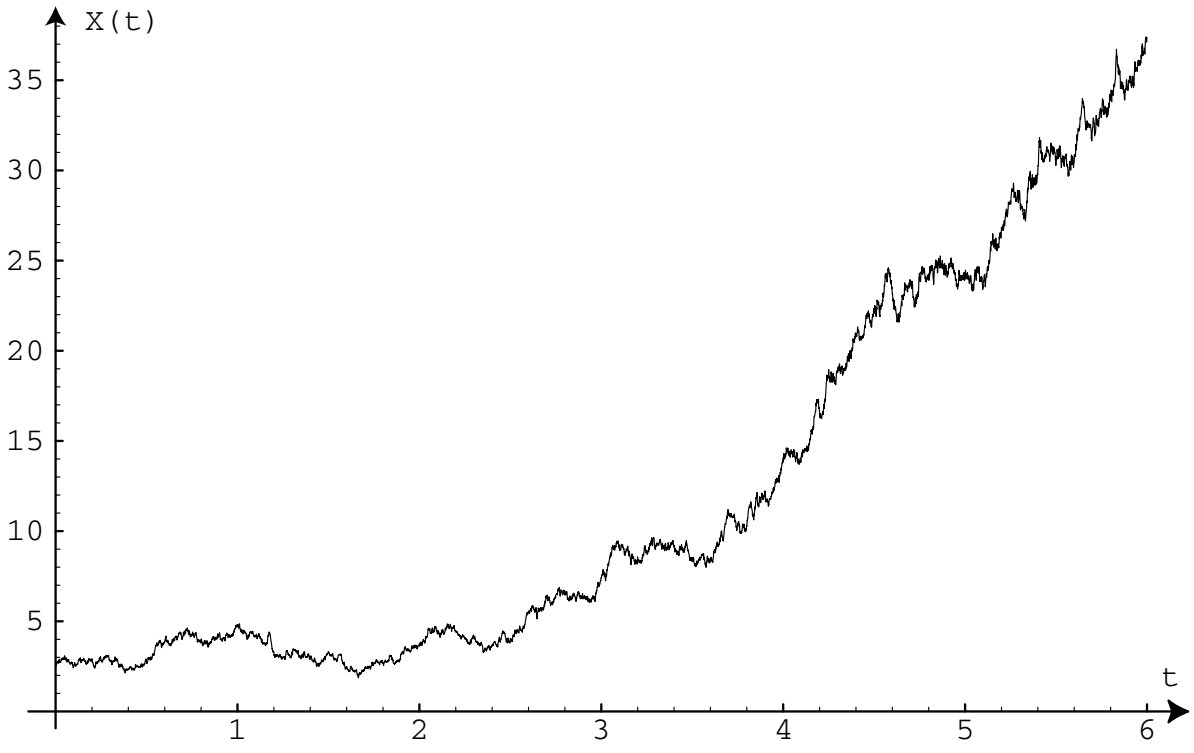

Figure 3.4: *A simulated path of the CIR process with* $x_0 = 2.5$, $a = 1$, $b = 0.5$, $\sigma = 1$.

Chapter 4

Parameter Estimation in CIR Processes

In Section 2.2.1 martingale estimating functions were discussed in detail. Theoretical results showed the asymptotic properties as there are consistency and asymptotic normality of the resulting estimators. Since these asymptotic properties do not depend on the length of the time interval between observations and since there are no constraints concerning parameter dependence, estimating parameters in diffusion processes via martingale estimating functions is an appropriate method in the case of discrete observations. In Chapter 3 properties and the behaviour of Cox–Ingersoll–Ross (CIR) processes (3.2) were described. Now we apply martingale estimating functions to simulated observations of CIR processes in order to investigate the numerical quality of martingale estimating functions. We note that in Bibby–Sørensen [7] martingale estimating functions were implemented and used to estimate parameters in the drift coefficient. Besides estimating parameters in the drift coefficient in Section 4.1 we also implemented martingale estimating functions that can be used to estimate parameters in the diffusion coefficient, see Section 4.3. We remark that in Fournie–Talay [32] and Overbeck–Rydén [78] different estimation methods are also applied to the CIR model and compared with the maximum likelihood method.

4.1 Estimation of Parameters in the Drift

Based on simulated observations of the CIR process (3.2) using the 1.5 strong Taylor scheme (see Kloeden–Platen [58], § 10.4) we estimate the parameters a and b in the drift via the discretized log-likelihood $\tilde{l}_n$ (2.18) and via martingale estimating functions $\tilde{G}_n$ based on the conditional mean (2.24). For both estimating functions $\tilde{l}_n$ and $\tilde{G}_n$ we obtain the estimators explicitly, see (4.1) and (4.2). We note that we also estimated a and b via the optimal martingale estimating function G_n^* (2.26) within the class (2.25). Since these estimates merely differ from the ones derived via $\tilde{G}_n$ and since the estimating equation cannot be solved explicitly but has to be solved numerically, here we concentrate on the approximate optimal estimating function $\tilde{G}_n$ (2.24). In Section 4.3 we give a more detailed discussion of using optimal martingale estimating functions. By means of the discretized log-likelihood $\tilde{l}_n$ we obtain for the estimators a_n^d and b_n^d

$$a_n^d = \frac{1}{n\Delta}\left(X_{n\Delta} - X_0 - \Delta\, b_n^d \sum_{i=1}^{n} X_{(i-1)\Delta}\right),$$

$$b_n^d = \frac{1}{\Delta}\frac{n\displaystyle\sum_{i=1}^{n}\frac{X_{i\Delta}}{X_{(i-1)\Delta}} - n^2 - (X_{n\Delta} - X_0)\displaystyle\sum_{i=1}^{n}\frac{1}{X_{(i-1)\Delta}}}{n^2 - \displaystyle\sum_{i=1}^{n} X_{(i-1)\Delta}\sum_{i=1}^{n}\frac{1}{X_{(i-1)\Delta}}}. \tag{4.1}$$

Using martingale estimating functions $\tilde{G}_n$, the formulae for the estimators $\tilde{a}_n$ and $\tilde{b}_n$ are

$$\tilde{a}_n = \frac{\tilde{b}_n}{1 - e^{\tilde{b}_n\Delta}} \cdot \frac{ne^{\tilde{b}_n\Delta} - \displaystyle\sum_{i=1}^{n}\frac{X_{i\Delta}}{X_{(i-1)\Delta}}}{\displaystyle\sum_{i=1}^{n}\frac{1}{X_{(i-1)\Delta}}},$$

$$\tilde{b}_n = \frac{1}{\Delta}\ln\left[\frac{n\displaystyle\sum_{i=1}^{n}\frac{X_{i\Delta}}{X_{(i-1)\Delta}} - \sum_{i=1}^{n} X_{i\Delta}\sum_{i=1}^{n}\frac{1}{X_{(i-1)\Delta}}}{n^2 - \displaystyle\sum_{i=1}^{n} X_{(i-1)\Delta}\sum_{i=1}^{n}\frac{1}{X_{(i-1)\Delta}}}\right]. \tag{4.2}$$

For the example with $a = 5$, $b = -1$, $\sigma = 1$ and $x_0 = 5$ equidistant observations are simulated. For each chosen number of observations n and step size Δ 500 simulations are considered. In Tables 4.1 and 4.2 the sample mean of the resulting estimates for a and b and the standard errors are listed for different values of Δ and n.

Δ	n	mean a^d	s.e. a^d	mean b^d	s.e. b^d
0.5	100	4.199	0.832	−0.845	0.169
0.5	500	3.995	0.371	−0.801	0.077
0.5	1000	3.953	0.258	−0.792	0.054
1.0	100	3.265	0.497	−0.655	0.100
1.0	500	3.180	0.227	−0.637	0.046
1.0	1000	3.162	0.155	−0.633	0.032

Table 4.1: *Mean and standard error of the estimator (a^d, b^d) based on the discretized log-likelihood $\tilde{l}_n$. Here $x_0 = 5$, $\sigma = 1$, and the true parameter values are $a = 5$ and $b = -1$. 500 simulations were made.*

Δ	n	mean $\tilde{a}$	s.e. $\tilde{a}$	mean $\tilde{b}$	s.e. $\tilde{b}$
0.5	100	5.567	1.523	−1.121	0.315
0.5	500	5.123	0.625	−1.027	0.130
0.5	1000	5.042	0.428	−1.010	0.089
1.0	100	5.565	1.717	−1.119	0.353
1.0	500	5.100	0.647	−1.022	0.133
1.0	1000	5.026	0.426	−1.006	0.087

Table 4.2: *Mean and standard error of the estimator $(\tilde{a}, \tilde{b})$ based on martingale estimating functions $\tilde{G}_n$. Here $x_0 = 5$, $\sigma = 1$, and the true parameter values are $a = 5$ and $b = -1$. 500 simulations were made.*

As we expect from the discussion in Section 2.2, p. 14 and p. 16, Tables 4.1 and 4.2 show that compared to the estimates based on $\tilde{l}_n$ the martingale estimating functions $\tilde{G}_n$ lead to better estimates. The estimates based on the martingale estimating function $\tilde{G}_n$ come closer to the true parameter value as n increases, and this independently of the

size of Δ. Hence, simulations confirm that martingale estimating functions are an appropriate parameter estimation method in the case of discrete observations. In the next subsection we consider estimating the diffusion parameter σ in a CIR process.

4.2 Simulation Study in the Nonergodic Case

As discussed in Chapter 3, CIR processes (3.2) are ergodic if and only if $b < 0$. Only for ergodic diffusions theoretical results on the asymptotic properties of estimators derived via martingale estimating functions are known. Therefore, based on a simulation study we are interested in the properties of the estimators for parameters in nonergodic CIR processes. As an example we consider the solution of

$$dX_t = (a + bX_t)\,dt + \sqrt{X_t}\,dW_t\,, \tag{4.3}$$

with $a = 2$, $b = 0.5$ and $X_0 = 5$ which is a nonergodic process, see Chapter 3, p. 31. We consider equidistant observations $X_0, X_\Delta, X_{2\Delta}, \ldots, X_{n\Delta}$ with $\Delta = 0.1$. The formulae for estimators a^d and b^d of the discretized log-likelihood function (2.18) and for the estimators $\tilde{a}$ and $\tilde{b}$ derived via $\tilde{G}_n$ (2.24) are known explicitly, see (4.1) and (4.2). Simulations of the process (4.3) are obtained using as an approximate simulation scheme the order 1.5 strong Taylor scheme, see (3.4).

In order to obtain reasonable simulated values at the time points Δ, 2Δ, $3\Delta, \ldots, n\Delta$, we split every interval of length Δ into N subintervals, and simulate the process at the endpoint of each subinterval, i.e. every Nth simulated point is taken as an observation. We remark that in the examples considered in the following, the results do not strongly depend on the chosen number N and therefore $N = 100$ or less seems to be reasonable. Every example is simulated 500 times and the sample mean of a^d, $\tilde{a}$, b^d and $\tilde{b}$ and the standard errors are shown in Table 4.3.

From Table 4.3 we deduce:

- The estimator $\tilde{b}$ obtained by the martingale estimating functions converges fast, seems to tend to the true parameter value and the standard error is very small.

n	m a^d	s.e. a^d	m $\tilde{a}$	s.e. $\tilde{a}$
100	2.5010	1.9884	2.4399	1.9401
200	2.4925	1.8485	2.4307	1.8027
300	2.3940	1.8677	2.3346	1.8214
400	2.4114	1.8307	2.3516	1.7852
500	2.5243	1.9716	2.4618	1.9227
600	2.3584	1.7307	2.2999	1.6878
700	2.3235	1.8737	2.2604	1.8240
800	2.5584	3.9723	2.2895	1.7263

n	m b^d	s.e. b^d	m $\tilde{b}$	s.e. $\tilde{b}$
100	0.5116	0.0070	0.4989	0.0233
200	0.5126	0.0019	0.4999	0.0017
300	0.5127	0.0005	0.5000	0.0001
400	0.5127	0.0002	0.5000	0.0000
500	0.5127	0.0000	0.5000	0.0000
600	0.5127	0.0000	0.5000	0.0000
700	0.5127	0.0000	0.5000	0.0000
800	0.5127	0.0000	0.5000	0.0000

Table 4.3: *Mean and standard error of estimates for a and b based on the discretized log-likelihood $\tilde{l}_n$ and on martingale estimating functions $\tilde{G}_n$.*

- The estimator b^d obtained by the discretized log-likelihood function converges fast, has a small bias (0.0127) and the standard error is very small.

- The estimates for a seem to tend to the right value 2 but they vary a lot, the standard error is large.

We remark that – as in the ergodic case – choosing a larger Δ has effects on a^d and b^d, but no effect on $\tilde{a}$ and $\tilde{b}$. As already mentioned, simulating more points between two observations merely influences the estimators for a and b. There is a strong dependence on the number of simulations and on the simulated paths. Note also that since the values of X_t increase exponentially, at some point they go beyond the

precision of the computer, and further estimation makes no sense (e.g. $X_{50} \approx 10^{12}$).

In order to understand why the estimators behave differently for a and b we consider the deterministic differential equation

$$dX_t = (a + bX_t)\,dt\,,\ a > 0,\, b > 0,\, X_0 = x_0\,,$$

with the solution

$$X_t = -\frac{a}{b} + (\frac{a}{b} + x_0)\,e^{bt}\,.$$

(X_t) increases exponentially with rate b and the parameter a influences the vertical shift. Adding the "stochastic part" $\sqrt{X_t}\,dW_t$, for large values of X_t the variations are small relative to X_t but absolutely large and therefore influence the estimation of a a lot. We remark that further simulation studies show that both diminishing Δ or treating b as known has essentially no effect on the estimates for a.

4.3 Estimation of Parameters in the Diffusion Coefficient

As discussed in Section 2.2.1, estimating parameters that only enter the diffusion coefficient is not possible by means of martingale estimating functions based on the conditional mean, but requires martingale estimating functions based on the conditional variance. We use the optimal martingale estimating functions K_n^* (2.36) within the class (2.35) and the approximately optimal martingale estimating functions $\tilde{K}_n$ (2.37) to obtain estimators for the parameters σ as well as for a and b in the CIR process (3.2). In Tables 4.4 and 4.5 the sample mean of the estimates and the standard error are listed for different values of n and Δ based on 1000 simulations using the 1.5 strong Taylor scheme (see Kloeden–Platen [58], § 10.4).

From Tables 4.4 and 4.5 we see that the behaviour of the estimator for σ is similar to the behaviour of the estimators for a and b: As n increases, the estimates come closer to the true parameter values, independently of the interval length Δ. Using K_n^* means a considerably higher computational effort than using $\tilde{K}_n$, since solving the estimating equation $K^*(a, b, \sigma) = 0$ has to be done numerically using Newton's

Δ	n	m a^*	s.e. a^*	m b^*	s.e. b^*	m ${\sigma^2}^*$	s.e. ${\sigma^2}^*$
0.5	100	5.5435	1.4724	-1.1142	0.3028	1.0256	0.1936
0.5	500	5.0998	0.6026	-1.0217	0.1240	1.0049	0.0834
0.5	1000	5.0558	0.4209	-1.0123	0.0875	1.0032	0.0605
1.0	100	5.4742	1.5501	-1.0984	0.3152	1.0389	0.2694
1.0	500	5.0751	0.6074	-1.0162	0.1232	1.0069	0.1071
1.0	1000	5.0347	0.4154	-1.0077	0.0849	1.0035	0.0766

Table 4.4: *Mean and standard error of the estimator* $(a^*, b^*, {\sigma^2}^*)$ *based on martingale estimating functions* K_n^**. Here* $x_0 = 5$*, and the true parameter values are* $a = 5$*,* $b = -1$ *and* $\sigma = 1$*. 1000 simulations were made.*

Δ	n	m $\tilde{a}$	s.e. $\tilde{a}$	m $\tilde{b}$	s.e. $\tilde{b}$	m $\tilde{\sigma}^2$	s.e. $\tilde{\sigma}^2$
0.5	100	5.5569	1.5084	-1.1197	0.3088	1.0273	0.1957
0.5	500	5.0999	0.6220	-1.0239	0.1278	1.0055	0.0862
0.5	1000	5.0585	0.4284	-1.0136	0.0890	1.0039	0.0615
1.0	100	5.4875	1.6476	-1.1009	0.3338	1.0458	0.2869
1.0	500	5.0783	0.6324	-1.0180	0.1282	1.0078	0.1133
1.0	1000	5.0354	0.4248	-1.0078	0.0869	1.0031	0.0790

Table 4.5: *Mean and standard error of the estimator* $(\tilde{a}, \tilde{b}, \tilde{\sigma}^2)$ *based on martingale estimating functions* $\tilde{K}_n$*. Here* $x_0 = 5$*, and the true parameter values are* $a = 5$*,* $b = -1$ *and* $\sigma = 1$*. 1000 simulations were made.*

method, whereas using $\tilde{K}_n$ leads to an explicit formula for $\tilde{\sigma}_n^2$ including $\tilde{a}_n$ and $\tilde{b}_n$.

$$\tilde{\sigma}_n^2 = \frac{\displaystyle\sum_{i=1}^{n} \left[X_{i\Delta} - \frac{1}{\tilde{b}_n} \left((\tilde{a}_n + \tilde{b}_n X_{(i-1)\Delta}) e^{\Delta \tilde{b}_n} - \tilde{a}_n \right) \right]^2 \frac{1}{X_{(i-1)\Delta}}}{\displaystyle\sum_{i=1}^{n} \frac{(\tilde{a}_n + 2\,\tilde{b}_n X_{(i-1)\Delta}) e^{2\Delta \tilde{b}_n} - 2\,(\tilde{a}_n + \tilde{b}_n X_{(i-1)\Delta}) e^{\Delta \tilde{b}_n} + \tilde{a}_n}{2 X_{(i-1)\Delta} \tilde{b}_n^2}}$$

Apart from the example in Tables 4.4 and 4.5 we simulated a lot of dif-

ferent CIR processes. In all cases both types K_n^* and $\tilde{K}_n$ of martingale estimating functions lead to similar estimates. Using K_n^* leads only to slightly better estimates with slightly smaller standard errors in comparison to $\tilde{K}_n$. Taking into account the enormous computational effort in using K_n^*, in our view this seems to give preference to using $\tilde{K}_n$ here.

Chapter 5

Some Generalizations of Bessel Processes

In many financial problems Bessel processes play an important role. They can be transformed via a time change to the exponential of Brownian motion with drift which is a standard model for stock prices (see (5.5)). This transformation is for instance used in the problem of pricing Asian options, see Geman–Yor [34]. Moreover, CIR processes discussed in Chapter 2 and 3 are time-space-transformed squared Bessel processes. By a time change, they can also be transformed to squared radial Ornstein–Uhlenbeck processes (for details see Section 5.1, I.5 and I.6). Motivated by this close relationship, in this chapter we obtain and discuss new results for radial Ornstein–Uhlenbeck and Bessel processes. This relation between Bessel processes and CIR processes is also used by Delbaen–Shirakawa [21, 22] and Szatzschneider [96].

The remarkable properties of Bessel processes are recalled in Section 5.1. In Section 5.2 we investigate deeply the first time radial Ornstein–Uhlenbeck processes hit a given barrier. An expression of the Laplace transform is given in Theorem 3. We introduce a more general class of Bessel processes in Section 5.3, Bessel processes with negative dimensions. Their interesting properties are discussed in detail and the results obtained are extended to a wider class of processes. As an extension of geometric Brownian motion, in Section 5.4 the exponential of Brownian

motion with inhomogeneous drift is investigated. Time reversal is an often used tool throughout this chapter. It is discussed in detail in the last section and moreover, new results are obtained.

5.1 Main Properties of Bessel Processes

We first give the definition of Bessel processes and state some of their properties. As for the study of Bessel processes we refer to Revuz–Yor [89] and Pitman–Yor [85, 87].

I.1 For every $\delta \geq 0$ and $x \geq 0$, the solution to the equation

$$X_t = x + \delta\, t + 2 \int_0^t \sqrt{|X_s|}\, dW_s$$

is unique and strong. In the case $\delta = 0, x = 0$, the solution X_t is identically zero and applying the comparison theorem (see Revuz–Yor [89] Theorem IX.(3.7)) we conclude $X_t \geq 0$ for all $\delta \geq 0$.

Definition 1 (BESQ$^\delta$) *For every $\delta \geq 0$ and $x \geq 0$ the unique strong solution to the equation*

$$X_t = x + \delta\, t + 2 \int_0^t \sqrt{X_s}\, dW_s$$

is called the square of a δ-dimensional Bessel process *started at x and is denoted by* BESQ$^\delta(x)$.

Denote the law of BESQ$^\delta(x)$ on $C(\mathbb{R}_+, \mathbb{R})$ by Q_x^δ. We call the number δ the *dimension* of BESQ$^\delta$. This notation arises from the fact that a BESQ$^\delta$ process X_t can be represented by the square of the Euclidean norm of δ-dimensional Brownian motion B_t: $X_t = |B_t|^2$. The number $\nu \equiv \delta/2 - 1$ is called the *index* of the process BESQ$^\delta$.

Definition 2 (BES$^\delta$) *The square root of* BESQ$^\delta(a^2)$, $\delta \geq 0$, $a \geq 0$ *is called the* Bessel process of dimension δ *started at a and is denoted by* BES$^\delta(a)$.

Denote the law of $\mathrm{BES}^\delta(a)$ by P_a^δ.

In the case $\delta \geq 2$ $\mathrm{BES}^\delta(a)$, $a > 0$, will never reach 0. For $\delta > 1$ a $\mathrm{BES}^\delta(a)$ process Z_t satisfies $E[\int_0^t (ds/Z_s)] < \infty$ and is the solution to the equation

$$Z_t = a + \frac{\delta - 1}{2} \int_0^t \frac{ds}{Z_s} + W_t \,.$$

For $\delta \leq 1$ the situation is less simple. For $\delta = 1$ we have with Itô-Tanaka's formula

$$Z_t = |W_t| = \tilde{W}_t + L_t \,,$$

where $\tilde{W}_t \equiv \int_0^t \mathrm{sgn}(W_s) dW_s$ is a standard Brownian motion, and L_t is the local time of Brownian motion. For a treatment of local times see e.g. Revuz–Yor [89], Chapter VI. For $\delta < 1$ we have

$$Z_t = a + \frac{\delta - 1}{2} \,\mathrm{p.v.} \int_0^t \frac{ds}{Z_s} + W_t \,, \tag{5.1}$$

where the principal value is defined as

$$\mathrm{p.v.} \int_0^t \frac{ds}{Z_s} \equiv \int_0^\infty x^{\delta-2} (L_t^x - L_t^0) dx$$

and the family of local times $(L_t^x, x \geq 0)$ is defined as

$$\int_0^t \varphi(Z_s) ds = \int_0^\infty \varphi(x) L_t^x x^{\delta-1} dx$$

for all Borel functions $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$, see Yor [103], Chapter 10. The decomposition (5.1) was obtained using the fact that a power of a Bessel process is another Bessel process time-changed, see (5.6).

I.2 Transition densities

Bessel (squared) processes are Markov processes and their transition densities are known explicitly. For $\delta > 0$, the transition density for BESQ^δ is equal to

$$q_t^\delta(x, y) = \frac{1}{2t} \left(\frac{y}{x}\right)^{\frac{\nu}{2}} \exp\left\{-\frac{x+y}{2t}\right\} I_\nu\left(\frac{\sqrt{xy}}{t}\right) , \tag{5.2}$$

where $t > 0$, $x > 0$, $\nu \equiv \frac{\delta}{2} - 1$ and I_ν is the modified Bessel function of the first kind of index ν. For $x = 0$ we have

$$q_t^\delta(0, y) = (2t)^{-\frac{\delta}{2}} \Gamma\left(\delta/2\right)^{-1} y^{\frac{\delta}{2}-1} \exp\left\{-\frac{y}{2t}\right\}. \tag{5.3}$$

For the case $\delta = 0$, the semi-group of BESQ^0 is equal to

$$Q_t^0(x, \cdot) = \exp\left\{-\frac{x}{2t}\right\} \varepsilon_0 + \tilde{Q}_t(x, \cdot), \tag{5.4}$$

where ε_0 is the Dirac measure in 0 and $\tilde{Q}_t(x, \cdot)$ has the density

$$q_t^0(x, y) = \frac{1}{2t} \left(\frac{y}{x}\right)^{-\frac{1}{2}} \exp\left\{-\frac{x+y}{2t}\right\} I_1\left(\frac{\sqrt{xy}}{t}\right).$$

The transition density for BES^δ is obtained from (5.2), (5.3) respectively (5.4) and is equal to

$$p_t^\delta(x, y) = \frac{1}{t} \left(\frac{y}{x}\right)^\nu y \exp\left\{-\frac{x^2+y^2}{2t}\right\} I_\nu\left(\frac{xy}{t}\right),$$

with $t > 0$, $x > 0$, and

$$p_t^\delta(0, y) = 2^{-\nu} t^{-(\nu+1)} \Gamma(\nu+1)^{-1} y^{2\nu+1} \exp\left\{-\frac{y^2}{2t}\right\},$$

and the semi-group for BES^0 is equal to

$$P_t^0(x, \cdot) = \exp\left\{-\frac{x^2}{2t}\right\} \varepsilon_0 + \tilde{P}_t(x, \cdot),$$

where ε_0 is the Dirac measure in 0 and $\tilde{P}_t(x, \cdot)$ has the density

$$p_t^0(x, y) = \frac{x}{t} \exp\left\{-\frac{x^2+y^2}{2t}\right\} I_1\left(\frac{xy}{t}\right).$$

I.3 Additivity property of Bessel squared processes

An important and well-known property of BESQ^δ processes with $\delta \geq 0$ is the following additivity property.

Theorem 2 (Shiga–Watanabe [95]) *For every $\delta, \delta' \geq 0$ and $x, x' \geq 0$:*

$$Q_x^\delta * Q_{x'}^{\delta'} = Q_{x+x'}^{\delta+\delta'},$$

where $Q_x^\delta * Q_{x'}^{\delta'}$ denotes the convolution of Q_x^δ and $Q_{x'}^{\delta'}$.

For a proof see Shiga–Watanabe [95] or Revuz–Yor[89], p. 410.

I.4 Scaling property and representations

BES^δ processes have the Brownian scaling property, i.e. if X is a $\mathrm{BES}^\delta(x)$, then the process $c^{-1}X_{c^2t}$ is a $\mathrm{BES}^\delta(x/c)$ for any $c>0$. BESQ^δ processes have the following scaling property: if X is a $\mathrm{BESQ}^\delta(x)$, then the process $c^{-1}X_{ct}$ is a $\mathrm{BES}^\delta(x/c)$.

An important result by Lamperti [64], see also Williams [99], is that the exponential of Brownian motion with drift can be represented as a time-changed Bessel process:

$$\exp(B_t+\nu t) = X^{(\nu)}\left(\int_0^t \exp 2(B_s+\nu s)\,ds\right), \quad t\geq 0\,, \tag{5.5}$$

where $(X^{(\nu)}(u), u\geq 0)$ is a Bessel process with index ν. We remark that because of the Brownian scaling property this result can be extended to a representation of $\exp(a\,B_t+\nu t)$. There is also a representation of $\exp(a\,B_t+\nu t)$ in terms of BESQ processes (see Geman–Yor [34]), and in terms of arbitrary powers of BES processes. This result is closely related to the fact that a power of a Bessel process is another Bessel process time-changed:

$$qZ_\nu^{1/q}(t) = Z_{\nu q}\left(\int_0^t \frac{ds}{Z_\nu^{2/p}(s)}\right), \tag{5.6}$$

where $\frac{1}{p}+\frac{1}{q}=1, \nu>-\frac{1}{q}$, see e.g. Revuz–Yor [89], Proposition 11.(1.11).

I.5 Transformation: CIR process – squared Bessel process

The Cox–Ingersoll–Ross (CIR) processes [16]

$$X_t = x+\int_0^t (a+bX_s)\,ds+\sigma\int_0^t \sqrt{X_s}\,dW_s\,, \tag{5.7}$$

for $t\geq 0$, $a\geq 0$, $\sigma>0$, were studied in detail in Chapter 3 and 4 including literature, behaviour and estimation in such processes. Here for a treatment of interest rate models including the CIR model we refer to Lamberton–Lapeyre [63], §6, and Musiela–Rutkowski [71]. We remark that these processes are also used in a population growth model, see Karlin–Taylor [53], p. 334.

The CIR process is a space-time transformed BESQ process, more explicitly: A $\mathrm{BESQ}^{\delta}(y)$ process Y can be transformed to the CIR process X by

$$X_t = e^{bt}\, Y\left(\frac{\sigma^2}{4b}(1-e^{-bt})\right), \tag{5.8}$$

where $\delta \equiv 4a/\sigma^2$.

I.6 Transformation: CIR process – squared radial Ornstein–Uhlenbeck process

A squared radial Ornstein–Uhlenbeck process is the solution to the equation

$$Y_t = y + \int_0^t (\delta - 2\lambda Y_s)\, ds + 2\int_0^t \sqrt{Y_s}\, dW_s$$

for $t \geq 0$, $\delta \geq 0$. Radial Ornstein–Uhlenbeck processes are studied in detail in Section 5.2, see (5.9). A squared radial Ornstein–Uhlenbeck process can be transformed by the time transformation $g(t) = \sigma^2 t/4$ to the CIR process X, where $\delta \equiv 4a/\sigma^2$, $\lambda \equiv -2b/\sigma^2$, with the notation in (5.7).

The transformations in I.5 and I.6 show the strong relation between CIR processes and squared Bessel respectively squared radial Ornstein–Uhlenbeck processes. In Section 5.2 we are considering first hitting times of squared radial Ornstein–Uhlenbeck processes, and by means of the transformation in I.6 we immediately see that, once the law of the first hitting time of a squared radial Ornstein–Uhlenbeck process is found, we know the law of the first hitting time of the CIR process.

5.2 First Hitting Times of Radial Ornstein–Uhlenbeck Processes

As motivated in the introduction (see I.6) we are interested in finding the law of first hitting times of squared radial Ornstein–Uhlenbeck processes. For a general discussion of first hitting times we refer to Breiman [11], Novikov [74, 75, 76], Shepp [94] and Yor [101].

First we recall the definition of (squared) radial Ornstein–Uhlenbeck processes. Let $\{W_t\}$ be a one-dimensional Brownian motion, $\lambda \in \mathbb{R}$,

$\delta \geq 0$ and $z \geq 0$. The solution to the equation

$$Z_t = z + \int_0^t (\delta - 2\lambda Z_s)\, ds + 2 \int_0^t \sqrt{|Z_s|}\, dW_s$$

is unique and strong (see Revuz–Yor [89] Chapter IX §3). Since in the case $\delta = 0, z = 0$, the solution is $Z_t \equiv 0$, we deduce from the comparison theorem (see Revuz–Yor [89] Theorem IX.(3.7)) $Z_t \geq 0$ for all $\delta \geq 0$, and hence the absolute value can be omitted; the solution of

$$Z_t = z + \int_0^t (\delta - 2\lambda Z_s)\, ds + 2 \int_0^t \sqrt{Z_s}\, dW_s \tag{5.9}$$

is called a *squared δ-dimensional radial Ornstein–Uhlenbeck process with parameter $-\lambda$* and its law on $C(\mathbb{R}_+, \mathbb{R})$ is denoted by ${}^{-\lambda}Q_z^\delta$. It is a Markov process; hence, the square root of this process is also a Markov process and is called a *δ-dimensional radial Ornstein–Uhlenbeck process with parameter $-\lambda$*. Denote its law on $C(\mathbb{R}_+, \mathbb{R})$ by ${}^{-\lambda}P_x^\delta$ where $x = \sqrt{z}$. For $\delta \geq 2$ and $z > 0$, it almost surely does not hit zero. For $\delta > 1$ it is the solution to the equation

$$dR_t = \left(\frac{\delta - 1}{2R_t} - \lambda R_t \right) dt + dW_t\,, \quad R_0 = x = \sqrt{z}\,.$$

For $\delta = 1$ we have with Itô-Tanaka's formula (see Revuz–Yor [89], Theorem 6.1.5)

$$|X_t| = |x| - \lambda \int_0^t |X_s|\, ds + \tilde{W}_t + L_t,$$

where X_t is a ${}^{-\lambda}P_x^1$ process, $\tilde{W} \equiv \int_0^t \operatorname{sgn}(W_s) dW_s$ is standard Brownian motion and L_t is the local time of Brownian motion. Using Girsanov's theorem (see Revuz–Yor [89], Chapter VIII, Theorem 8.1.4) we obtain the relationship

$${}^{-\lambda}P_x^\delta\big|_{\mathcal{F}_t} = \exp\left\{ -\lambda \int_0^t R_s dW_s - \frac{\lambda^2}{2} \int_0^t R_s^2 ds \right\} \cdot P_x^\delta\big|_{\mathcal{F}_t}\,,$$

where P_x^δ denotes the law of a $\mathrm{BES}^\delta(x)$ process. Hence for $\delta < 1$ we obtain from (5.1)

$$R_t = x + \frac{\delta - 1}{2}\, \text{p.v.} \int_0^t \frac{ds}{R_s} - \lambda \int_0^t R_s ds + \hat{W}_t\,,$$

where $\hat{W}_t$ is standard Brownian motion under ${}^{-\lambda}P^{\delta}$.

Our aim is to find the law of

$$T_y = \inf\{t \,|\, R_t = y\}\,, \tag{5.10}$$

the first time a radial Ornstein–Uhlenbeck process $\{R_t\}$ with parameter $-\lambda$ starting in $x > 0$ hits the level y, $0 < y < x$. For $\delta < 2$, we have ${}^{-\lambda}P_x^{\delta}(T_0 < \infty) > 0$, that is, the process $\{R_t\}$ may reach 0; if $\delta < 2$ and $\lambda > 0$, then ${}^{-\lambda}P_x^{\delta}(T_0 < \infty) = 1$, that is, $\{R_t\}$ reaches 0 a.s. and hence every y a.s., $0 < y < x$.

We remark, that since:

$$T_y = \inf\{t \,|\, R_t = y\} = \inf\{t \,|\, R_t^2 = y^2\}\,,$$

once we found the law of (5.10) we have solved our original problem: finding the law of the first time a squared radial Ornstein–Uhlenbeck process with $\delta < 2$ starting in $a\,(= x^2)$ hits $b\,(= y^2)$, $0 < b < a$.

Call ${}^{-\lambda}P_x^{\delta}$ the law of a δ-dimensional radial Ornstein–Uhlenbeck process with parameter $-\lambda$. The density ${}^{-\lambda}p_x^{\delta}(t)$ of the first hitting time of 0 of a radial Ornstein–Uhlenbeck process is calculated in Elworthy–Li–Yor [26], Corollary 3.10, by using a time reversal argument, that is, for $y = 0$ the problem is solved:

$${}^{-\lambda}p_x^{\delta}(t) = \frac{x^{2-\delta}}{2^{\nu}\Gamma(\nu)}\, \exp\left[\tfrac{\lambda}{2}(\delta t + x^2(1 - \coth(\lambda t)))\right] \left[\frac{\lambda}{\sinh(\lambda t)}\right]^{\frac{4-\delta}{2}}, \tag{5.11}$$

where $\delta < 2$, $\lambda > 0$, $x > 0$ and $\nu = \frac{4-\delta}{2} - 1$.

In the following, let $T_{x\to y}$ denote the first time a radial Ornstein–Uhlenbeck process with dimension $\delta < 2$, starting in $x > 0$, hits y, $0 \le y < x$. Why calculating the density of $T_{x\to y}$ is more complicated than calculating the density of $T_{x\to 0}$ will be explained later. We have

$$T_{x\to 0} \overset{(\text{law})}{=} T_{x\to y} + T_{y\to 0}\,,$$

where $T_{x\to y}$ and $T_{y\to 0}$ are independent because of the strong Markov property. Hence, we have for the Laplace transforms (LTs)

$${}^{-\lambda}E_x^{\delta}\left[\exp\left(-\mu T_{x\to 0}\right)\right] = {}^{-\lambda}E_x^{\delta}\left[\exp\left(-\mu T_{x\to y}\right)\right]\, {}^{-\lambda}E_y^{\delta}\left[\exp\left(-\mu T_{y\to 0}\right)\right]\,,$$

or equivalently,

$$ {}^{-\lambda}E_x^{\delta}\left[\exp\left(-\mu T_{x\to y}\right)\right] = \frac{\phi_x(\mu)}{\phi_y(\mu)}\,, \tag{5.12} $$

where

$$ \phi_x(\mu) = {}^{-\lambda}E_x^{\delta}\left[\exp\left(-\mu T_{x\to 0}\right)\right] = \int_0^{\infty} \exp(-\mu t)\ {}^{-\lambda}p_x^{\delta}(t)\,dt. $$

The rest of this section is divided into two parts. First, we consider the case $\lambda = 0$, i.e. Bessel processes, where we can find explicit expressions of the LT in (5.12) for some examples. In the second part we consider more generally radial Ornstein–Uhlenbeck processes and derive as the main result an expression of the LT of first hitting times, see Theorem 3, by following an idea we got by studying Breiman [11].

First hitting times of Bessel processes

In order to find an explicit expression of the LT of first hitting times of Bessel processes, it is convenient to consider Bessel processes time-reversed. We refer to Section 5.5 where time reversal is studied in detail. Let $\{R_t\}$ be a Bessel process with dimension $\delta < 2$. The time reversed process

$$ \hat{R}_u \equiv R_{T_0-u}\,, \quad u \geq 0, $$

is a $\hat{\delta}$-dimensional Bessel process starting in 0, with $\hat{\delta} \equiv (4-\delta)$. The process $(\hat{R}_u)$ is transient since $\hat{\delta} > 2$. We know by the time reversal Theorem 6 that:

$$ \left(R_{T_0-u},\ u \leq T_0\right) \overset{(\text{law})}{=} \left(\hat{R}_u,\ u \leq \hat{L}_x\right)\,, \tag{5.13} $$

where

$$ \hat{L}_x = \sup\{t \,|\, \hat{R}_t = x\} $$

is the last exit time of x by the process $\{\hat{R}_t\}$. In particular, as remarked in Getoor–Sharpe [37], Sharpe [93], $\hat{L}_x$ under $P_0^{\hat{\delta}}$ has the same law as T_0 under P_x^{δ}. We know (see Getoor [36], Pitman–Yor [85]):

$$ \hat{L}_x \overset{(\text{law})}{=} \frac{x^2}{2Z_{\hat{\nu}}}\,, \tag{5.14} $$

where $Z_{\hat{\nu}}$ is a Gamma variable with parameter $\hat{\nu} \equiv \frac{\hat{\delta}}{2} - 1$, i.e.:

$$P(Z_{\hat{\nu}} \in dt) = \frac{t^{\hat{\nu}-1}e^{-t}}{\Gamma(\hat{\nu})}\, dt\,, \tag{5.15}$$

and hence

$$P_x^{\delta}(T_0 \in dt) = P_0^{\hat{\delta}}(\hat{L}_x \in dt) = \frac{1}{t\Gamma(\hat{\nu})}\left(\frac{x^2}{2t}\right)^{\hat{\nu}} e^{-\frac{x^2}{2t}}\, dt\,. \tag{5.16}$$

Let $\hat{L}_{0\to x}$ respectively $\hat{L}_{0\to y}$ denote the last exit time of x respectively y by a $\hat{\delta}$-dimensional Bessel process $(\hat{R}_u)$ starting in 0. With $\hat{L}_{y\nearrow x}$ we denote the last exit time of x by $(\hat{R}_u)$ starting in y at the last exit time of y, i.e., we consider the process starting in y but never visiting y again; distinguish carefully between $\hat{L}_{y\nearrow x}$ and $\hat{L}_{y\to x}$, the last exit time of x by $(\hat{R}_u)$ starting in y. Then we have

$$\hat{L}_{0\to x} \overset{(\text{law})}{=} \hat{L}_{0\to y} + \hat{L}_{y\nearrow x}\,,$$

where $\hat{L}_{0\to y}$ and $\hat{L}_{y\nearrow x}$ are independent because of the strong Markov property of last exit times (for a survey see Millar [70]). Since $\hat{L}_{y\nearrow x}$ has the same law as $T_{x\to y}$, our aim is to find an explicit expression of the LT of $\hat{L}_{y\nearrow x}$:

$$E_0^{\hat{\delta}}\left[\exp\left(-\tfrac{\alpha^2}{2}\hat{L}_{y\nearrow x}\right)\right] = \frac{E_0^{\hat{\delta}}\left[\exp\left(-\tfrac{\alpha^2}{2}\hat{L}_{0\to x}\right)\right]}{E_0^{\hat{\delta}}\left[\exp\left(-\tfrac{\alpha^2}{2}\hat{L}_{0\to y}\right)\right]}\,.$$

Now we can explain the difference between calculating the density of $T_{x\to 0}$ and the density of $T_{x\to y}$ for $y \neq 0$ of a Bessel process. The density of $T_{x\to 0}$ is obtained by considering the time reversed process started in 0 and using (5.14) and that $T_{x\to 0} \overset{(\text{law})}{=} \hat{L}_{0\to x}$. The density of $T_{x\to 0}$ of a radial Ornstein–Uhlenbeck process (5.11) was obtained in a similar way, additionally using a Girsanov transformation. As for the density of $T_{x\to y}$ we cannot use the same time reversal argument since we are interested in the time reversed process *started at the last exit time of* y, and this is the time reversed process started in y conditioned on never hitting y again. Figure 5.1 illustrates the situation. With (5.16) and with an integral representation of $K_{\hat{\nu}}$, i.e. the modified Bessel function

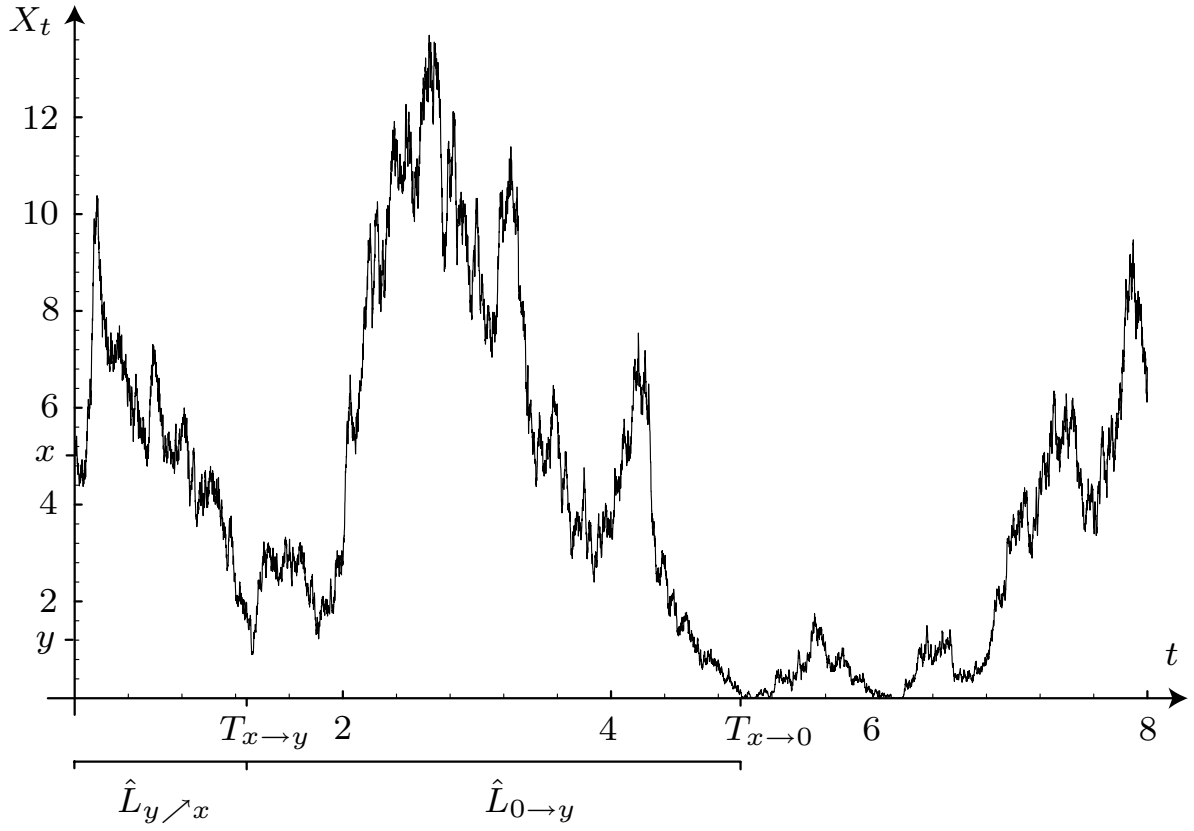

Figure 5.1: *A simulated path of a* BES^δ *process starting in* 5.

of the second kind of order $\hat{\nu}$, (see e.g. Lebedev [65] p. 119, (5.10.25)) we obtain the LT of $\hat{L}_{0\to y}$:

$$E_0^{\hat{\delta}}\left[\exp\left(-\tfrac{\alpha^2}{2}\hat{L}_{0\to y}\right)\right] = \frac{(\alpha y)^{\hat{\nu}} K_{\hat{\nu}}(\alpha\, y)}{2^{\hat{\nu}-1}\Gamma(\hat{\nu})}. \tag{5.17}$$

Thus (see also Getoor [36], Getoor–Sharpe [37], Kent [54]):

$$E_x^{\delta}\left[\exp\left(-\tfrac{\alpha^2}{2}T_y\right)\right] = E_0^{\hat{\delta}}\left[\exp\left(-\tfrac{\alpha^2}{2}\hat{L}_{y\nearrow x}\right)\right] = \left(\frac{x}{y}\right)^{\hat{\nu}} \frac{K_{\hat{\nu}}(\alpha\, x)}{K_{\hat{\nu}}(\alpha\, y)}\,. \tag{5.18}$$

Now we want to find a tractable form of $K_{\hat{\nu}}(\alpha\, x)/K_{\hat{\nu}}(\alpha\, y)$ to obtain the density explicitly. We cannot solve this problem in general but we have found explicit forms in some special cases.

a) First, we consider Brownian motion, that is, the case $\delta = 1$ (equiv-

alently $\hat{\delta} = 3$ or $\hat{\nu} = \frac{1}{2}$). In this case we have

$$\hat{L}_{0\to x} \overset{\text{(law)}}{=} T_{x\to 0} \overset{\text{(law)}}{=} T_{y\to 0} + (\tilde{T}_{(x-y)\to 0}),$$

where $(\tilde{T}_{(x-y)\to 0})$ denotes an independent copy of $T_{x\to y}$. Moreover, we know the density of $T_{x\to 0}$ for Brownian motion:

$$q_x(t) = \frac{x}{\sqrt{2\pi t^3}}\, e^{-\frac{x^2}{2t}}, \tag{5.19}$$

that is, we know the density of $T_{(x-y)\to 0}$ and therefore the density of $\hat{L}_{y\nearrow x}$. Note that

$$E_x\left[\exp\left(-\tfrac{\alpha^2}{2} T_{x\to 0}\right)\right] = \int_0^\infty e^{-\frac{\alpha^2 t}{2}} q_x(t)\, dt = e^{-\alpha x}. \tag{5.20}$$

b) In the cases $\delta = -1, -3, -5, \ldots$, or equivalently $\hat{\nu} = \frac{3}{2}, \frac{5}{2}, \frac{7}{2}, \ldots$, we need to look precisely at

$$\frac{K_{n+\frac{1}{2}}(\alpha\, x)}{K_{n+\frac{1}{2}}(\alpha\, y)}, \quad n \in \mathbb{N}.$$

We know that

$$K_{n+\frac{1}{2}}(z) = \sqrt{\frac{\pi}{2}} P_n(z) e^{-z} z^{-\left(n+\frac{1}{2}\right)}, \tag{5.21}$$

where

$$P_n(z) = \sum_{k=0}^{n} \frac{(n+k)!}{(n-k)!}\, \frac{z^{n-k}}{2^k k!}, \quad n \in \mathbb{N}_0,$$

see Ismail–Kelker [48], p. 82. From (5.21) we deduce

$$\frac{K_{n+\frac{1}{2}}(\alpha\, x)}{K_{n+\frac{1}{2}}(\alpha\, y)} = \left(\frac{y}{x}\right)^{n+\frac{1}{2}} e^{-\alpha(x-y)} \frac{P_n(\alpha\, x)}{P_n(\alpha\, y)}.$$

Hence, the LT of $\hat{L}_{y\nearrow x}$ in $\frac{\alpha^2}{2}$ is

$$E_0^{\hat{\delta}}\left[\exp\left(-\tfrac{\alpha^2}{2}\hat{L}_{y\nearrow x}\right)\right] = \left(\frac{x}{y}\right)^{\hat{\nu}} \frac{K_{\hat{\nu}}(\alpha\, x)}{K_{\hat{\nu}}(\alpha\, y)} = e^{-\alpha(x-y)} \frac{P_n(\alpha\, x)}{P_n(\alpha\, y)} \tag{5.22}$$

in the case $\hat{\nu} = n + \frac{1}{2}$.

Let us look at the case $n = 1$. For the right term in (5.22) we have the following equalities:

$$e^{-\alpha(x-y)}\,\frac{P_1(\alpha\,x)}{P_1(\alpha\,y)} = e^{-\alpha(x-y)}\,\frac{\alpha x+1}{\alpha y+1} \tag{5.23}$$

$$= \frac{\alpha x}{\alpha y+1}\,e^{-\alpha(x-y)} + \frac{1}{\alpha y+1}\,e^{-\alpha(x-y)} \tag{5.24}$$

$$= \frac{x}{y}\left(\frac{(\alpha y+1)-1}{\alpha y+1}\right)e^{-\alpha(x-y)} + \frac{e^{-\alpha(x-y)}}{\alpha y+1} \tag{5.25}$$

$$= \underbrace{\tfrac{x}{y}\,e^{-\alpha(x-y)}}_{A} + \underbrace{\left(1-\tfrac{x}{y}\right)\frac{1}{\alpha y+1}\,e^{-\alpha(x-y)}}_{B}\,. \tag{5.26}$$

We know that A is the LT in $\frac{\alpha^2}{2}$ of $\frac{x}{y}\,q_{x-y}(t)$, see (5.19) and (5.20). As for the second term B we have

$$B = \left(1-\tfrac{x}{y}\right)e^{-\alpha(x-y)}\int_0^\infty \exp[-(1+\alpha y)\,u]\,du \tag{5.27}$$

$$= \left(1-\tfrac{x}{y}\right)\int_0^\infty e^{-u}e^{-\alpha(yu+x-y)}\,du, \tag{5.28}$$

which is the LT in $\frac{\alpha^2}{2}$ of

$$\left(1-\tfrac{x}{y}\right)\int_0^\infty e^{-u}\,q_{(yu+x-y)}(t)\,du.$$

And since

$$\hat{q}(t) \equiv \tfrac{x}{y}\,q_{x-y}(t) + \left(1-\tfrac{x}{y}\right)\int_0^\infty e^{-u}\,q_{(yu+x-y)}(t)\,du$$

is a convex combination of densities we have that

$$e^{-\alpha(x-y)}\,\frac{P_1(\alpha x)}{P_1(\alpha y)}$$

is the LT in $\frac{\alpha^2}{2}$ of $\hat{q}(t)$.

First hitting times of radial Ornstein–Uhlenbeck processes

Our aim is to find an explicit expression of the LT (5.12) of $T_{x\to y}$ of a δ-dimensional radial Ornstein–Uhlenbeck process R with $\delta < 2$. But

before we derive the main result Theorem 3 of this section, we consider the time reversed radial Ornstein–Uhlenbeck process $\hat{R}$ started in 0. We can write the process after $\hat{L}_{0\to y}$, the last exit of y, as a diffusion, that is, as the solution of a stochastic differential equation. This we obtain by using a specific technique, the "enlargement of filtration". For a treatment see e.g. Jeulin [51] and Yor [103] §12. Heuristically spoken, we enlarge the original filtration progressively, so that the last exit time $\hat{L}_{0\to y}$ becomes a stopping time. Applying Theorem 12.4 in Yor [103] we obtain

Proposition 1

$$\tilde{R} \equiv \hat{R}_{(\hat{L}_{0\to y}+u)} = y + \int_0^u b(\tilde{R}_v)dv - \int_0^u \frac{s'(\tilde{R}_v)}{s(y) - s(\tilde{R}_v)}\, 1_{(\tilde{R}_v > y)} dv + \tilde{W}_u\,, \tag{5.29}$$

where $u \geq 0$, b is the drift and s is the scale function of the transient diffusion $\hat{R}$.

Thus we have written the process after $\hat{L}_{0\to y}$, that is, the time reversed process started in y conditioned on never hitting y again, as a diffusion $\tilde{R}$. As an illustration consider the process $\hat{R}$ to be a transient BES process $\hat{X}$, i.e. a BES process with index $\nu > 0$, started in 0. Its scale function may be chosen as $s(x) = -x^{-2\nu}$ and we obtain from (5.29)

$$\tilde{X}_u \equiv \hat{X}_{(\hat{L}_{0\to y}+u)} = y + \int_0^u \frac{1}{\tilde{X}_v} \frac{(\nu + \frac{1}{2})\tilde{X}_v^{2\nu} + (\nu - \frac{1}{2})y^{2\nu}}{\tilde{X}_v^{2\nu} - y^{2\nu}}\, dv + \tilde{W}_u. \tag{5.30}$$

For a BES$^3(0)$ process $\hat{X}$ (5.30) reduces to

$$\tilde{X}_u \equiv \hat{X}_{(\hat{L}_{0\to y}+u)} = y + \int_0^u \frac{dv}{\tilde{X}_v - y} + \tilde{W}_u.$$

Hence for a BES$^3(0)$ process X we have

$$(X_{L_y+u} - y,\, u \geq 0) \overset{(\text{law})}{=} (X_u,\, u \geq 0)\,.$$

In general, the transition density $\tilde{p}$ of the diffusion $\tilde{R}$ in (5.29) is unknown. Note that if it were known, we would obtain the density of the last hitting time of x by the process $\tilde{R}$ immediately from the formula

$$P_y(\tilde{L}_x \in dt) = -\frac{1}{2s(x)}\, \tilde{p}_t(y,x)\, dt\,, \tag{5.31}$$

see Borodin–Salminen [10] IV.43, Revuz–Yor [89] VII.(4.16), where $\tilde{R}_0 = y$ and s is the scale function of $\tilde{R}$ with $\lim_{a\downarrow 0} s(a) = -\infty$ and $s(\infty) = 0$.

Finally, let us now derive the main result of this section by pursuing an idea concerning first hitting times of radial Ornstein–Uhlenbeck processes we got by studying Breiman [11]. Breiman considers an Ornstein–Uhlenbeck process $Y(u) = e^{-u}B(e^{2u})$ with B standard Brownian motion, $B(1) = 0$, and gives the LT of the first hitting time T_c for Y from the boundaries $\pm c$

$$E(e^{-\alpha T_c}) = \tilde{D}_\alpha^{-1}(c)\,,$$

where

$$\tilde{D}_\alpha(c) = \frac{2^{1-\lambda/2}}{\Gamma(\frac{\lambda}{2})} \int_0^\infty e^{-\frac{t^2}{2}} t^{\lambda-1} \cosh(ct)\, dt\,.$$

A δ-dimensional radial Ornstein–Uhlenbeck process R_t with parameter $-\lambda$, $R_0 = 0$, can be written as

$$R_t = e^{-\lambda t} X\left(\frac{e^{2\lambda t} - 1}{2\lambda}\right),$$

where X is a BES^δ process, $X_0 = 0$. We are interested in the first time T_c, R_t hits the level $c > 0$.

Analogously to Breiman [11] let us assume that D_α is a function such that $\{D_\alpha(R_t)e^{-\alpha t},\, t \geq 0\}$ is a martingale with respect to the filtration of $X\left(\frac{e^{2\lambda t}-1}{2\lambda}\right)$. Equivalently, $\{D_\alpha(\frac{1}{\sqrt{2\lambda u+1}}X_u)(2\lambda u+1)^{-\frac{\alpha}{2\lambda}},\, u \geq 0\}$ is a martingale with respect to the filtration of X_u, which means that $\{D_\alpha(\frac{r}{\sqrt{2\lambda u+1}})(2\lambda u+1)^{-\frac{\alpha}{2\lambda}},\, r \geq 0,\, u \geq 0\}$ is a space-time harmonic function with respect to $\frac{1}{2}\frac{d^2}{dr^2} + \frac{\delta-1}{2r}\frac{d}{dr} + \frac{d}{du}$.

Applying the optional stopping theorem (see Revuz–Yor [89] §2.3, p. 65) we have

$${}^{-\lambda}E^\delta[D_\alpha(R_{T_c})e^{-\alpha T_c}] = D_\alpha(0)\,,$$

thus

$${}^{-\lambda}E^\delta[e^{-\alpha T_c}] = \frac{D_\alpha(0)}{D_\alpha(c)}. \tag{5.32}$$

Now we are motivated to find the function D_α explicitly. First, we derive the space-time harmonic functions for Bessel processes, with integer dimensions.

We know by Widder [98], Theorem 14.1.1, that the most general $(h(x,t), x \in \mathbb{R}^n, t \geq 0)$ space-time harmonic function with respect to $\frac{1}{2}\Delta x + \frac{d}{dt}$ is of the form

$$h(x,t) = \int_{\mathbb{R}^n} \exp\left(\xi \cdot x - \frac{|\xi|^2 t}{2}\right) m(d\xi), \tag{5.33}$$

where m is of bounded variation and nondecreasing. Consider $h(B_t, t)$ with B_t an n-dimensional Brownian motion with $B_0 = 0$. From (5.33) we obtain the general space-time harmonic function $(\tilde{h}(r,t),\ r,\ t \in \mathbb{R}_+)$ with respect to $\frac{1}{2}\frac{d^2}{dr^2} + \frac{n-1}{2r}\frac{d}{dr} + \frac{d}{dt}$ via

$$\begin{aligned} \tilde{h}(r,t) &\equiv E_0\big(h(B_t,t)|\,|B_t| = r\big) = \int_{S^{n-1}} h(\theta r, t)\,\sigma(d\theta) \\ &= \int_{\mathbb{R}^n}\int_{S^{n-1}} \exp(\xi\cdot\theta r - \frac{|\xi|^2 t}{2})\,\sigma(d\theta)\,m(d\xi), \end{aligned} \tag{5.34}$$

where σ is the uniform probability measure on the unit sphere S^{n-1}. We know the following integral representation for the Bessel function I_ν from formula *(5.c2)* in Pitman–Yor [85]

$$I_\nu(|\eta|) = \frac{|\frac{\eta}{2}|^\nu}{\Gamma(\nu+1)} \int_{S^{n-1}} \exp(\eta\cdot\theta)\,\sigma(d\theta)\,,$$

$\eta \in \mathbb{R}^n$, $\nu = \frac{n}{2} - 1$, and inserting in (5.34) we have

$$\tilde{h}(r,t) = \Gamma(\nu+1)\left(\frac{2}{r}\right)^\nu \int_{\mathbb{R}^n} I_\nu(r|\xi|)\exp\left(-\frac{|\xi|^2 t}{2}\right)\frac{1}{|\xi|^\nu}\,m(d\xi)\,.$$

Substituting $|\xi|$ by $u \in \mathbb{R}_+$ we obtain the space time harmonic function $\tilde{h}$ with respect to a Bessel process

$$\tilde{h}(r,t) = \Gamma(\nu+1)\left(\frac{2}{r}\right)^\nu \int_0^\infty I_\nu(ru) e^{-\frac{u^2 t}{2}} u^{-\nu}\,\tilde{m}(du)\,.$$

Now we want to find a function D_α such that

$$\begin{aligned} D_\alpha\left(\frac{r}{\sqrt{2\lambda t+1}}\right)(2\lambda t+1)^{-\frac{\alpha}{2\lambda}} &= \tilde{h}(r,t) \\ &= \Gamma(\nu+1)\left(\frac{2}{r}\right)^\nu \int_0^\infty I_\nu(ru)\,e^{-\frac{u^2 t}{2}} f(u)\,du\,, \end{aligned}$$

with a L^1-function $f : \mathbb{R}_+ \to \mathbb{R}_+$.

Let us first consider

$$D_\alpha\left(\frac{r}{\sqrt{2\lambda t}}\right)(2\lambda t)^{-\frac{\alpha}{2\lambda}} = \Gamma(\nu+1)\left(\frac{2}{r}\right)^\nu \int_0^\infty I_\nu(ru)\, e^{-\frac{u^2 t}{2}}\, f(u)\, du\,,$$

equivalently,

$$D_\alpha(y) = (2\lambda t)^{\frac{\alpha}{2\lambda}}\,\Gamma(\nu+1)\left(\frac{2}{\sqrt{2\lambda t}\, y}\right)^\nu \int_0^\infty I_\nu(\sqrt{2\lambda t}\, yu)\, e^{-\frac{u^2 t}{2}}\, f(u)\, du$$

$$= (2\lambda t)^{\frac{\alpha-\lambda}{2\lambda}}\,\Gamma(\nu+1)\left(\frac{2}{\sqrt{2\lambda t}\, y}\right)^\nu \int_0^\infty I_\nu(\eta y)\, e^{-\frac{\eta^2}{4\lambda}}\, f\left(\frac{\eta}{\sqrt{2\lambda t}}\right) d\eta. \quad (5.35)$$

The right hand side of (5.35) should not depend on t, hence

$$f(x) \equiv c_0\, x^{\frac{\alpha-\lambda}{\lambda}-\nu}, \quad (5.36)$$

with a constant $c_0 > 0$. We see that $D_\alpha(\frac{r}{\sqrt{2\lambda t}})(2\lambda t)^{-\frac{\alpha}{2\lambda}} = \tilde{h}(r,t)$, and certainly also $\tilde{h}(r, t+c)$ with a constant c, is a space-time harmonic function with respect to $\frac{1}{2}\frac{d^2}{dr^2} + \frac{\delta-1}{2r}\frac{d}{dr} + \frac{d}{dt}$. Thus

$$D_\alpha\left(\frac{r}{\sqrt{2\lambda t+1}}\right)(2\lambda t+1)^{-\frac{\alpha}{2\lambda}}$$

$$= \Gamma(\nu+1)\left(\frac{2}{r}\right)^\nu \int_0^\infty I_\nu(ru)\, e^{-\frac{u^2}{2}\left(t+\frac{1}{2\lambda}\right)} f(u)\, du$$

with f in (5.36) and we find

$$D_\alpha(y) = \Gamma(\nu+1)\left(\frac{2}{y}\right)^\nu \int_0^\infty I_\nu(\eta y)\, e^{-\frac{\eta^2}{4\lambda}}\, \eta^{\frac{\alpha-\lambda}{\lambda}-\nu}\, d\eta.$$

With (5.83) we see

$$D_\alpha(0) = \int_0^\infty e^{-\frac{\eta^2}{4\lambda}}\, \eta^{\frac{\alpha-\lambda}{\lambda}} d\eta = 2^{\frac{a}{\lambda}-1}\, \lambda^{\frac{a}{2\lambda}}\, \Gamma(\tfrac{\alpha}{2\lambda})\,,$$

and finally,

Theorem 3 *The LT of the first time $T_c = \inf\{t \,|\, R_t = c\}$ a δ-dimensional radial Ornstein–Uhlenbeck process R_t with $R_0 = 0$ and with parameter $-\lambda$ hits the level c is*

$${}^{-\lambda}E^\delta(e^{-\alpha T_c}) = \frac{2^{-1-\nu+\frac{\alpha}{\lambda}}\, c^\nu\, \Gamma(\frac{\alpha}{2\lambda})\, \lambda^{\frac{\alpha}{2\lambda}}}{\Gamma(\nu+1)\displaystyle\int_0^\infty I_\nu(\eta c)\, e^{-\frac{\eta^2}{4\lambda}}\, \eta^{\frac{\alpha-\lambda}{\lambda}-\nu} d\eta}. \quad (5.37)$$

5.3 BESQ Processes with Negative Dimensions and Extensions

Bessel processes with nonnegative dimension $\delta \geq 0$ and starting point $x \geq 0$ are well-studied, see Section 5.1, I.1-I.4 and e.g. Revuz–Yor [89], Chapter XI. It seems to be quite natural also to consider Bessel processes with nonnegative dimension and negative starting point. As will be developed later this case is strongly related to Bessel processes with negative dimensions. Therefore we are motivated to extend the definition of BESQ_x^δ processes (see I.1) and to introduce the class of BESQ_x^δ processes with arbitrary δ, $x \in \mathbb{R}$. We will derive and discuss the properties of this class of processes. Moreover, as an extension we generalize squared δ-dimensional radial Ornstein–Uhlenbeck processes.

Definition 3 *The solution to the stochastic differential equation*

$$dX_t = \delta\, dt + 2\sqrt{|X_t|}\, dW_t\,, \quad X_0 = x\,, \tag{5.38}$$

where $\{W_t\}$ is a one-dimensional Brownian motion, $\delta \in \mathbb{R}$ and $x \in \mathbb{R}$, is called the square of a δ-dimensional Bessel process*, starting in x, and is denoted by* BESQ_x^δ.

Equation (5.38) has a unique strong solution (see Revuz–Yor [89], Chapter IX §3). Denote its law by Q_x^δ. First, we want to investigate the behaviour of a BESQ_x^δ process, starting in $x > 0$ with dimension $\delta \leq 0$. In the case $\delta = 0$ the process reaches 0 in finite time and stays there. As for the case $\delta < 0$, we deduce from the comparison theorem (see Revuz–Yor [89] Theorem IX.(3.7)) that this process is smaller than the process with $\delta = 0$, hence, that 0 is also reached in finite time. Let us consider the behaviour of a BESQ_x^δ process $\{X_t\}$ with $\delta < 0$ and $x > 0$ after it reached 0; we find:

$$\tilde{X}_u \equiv X_{T_0+u} = \delta\, u + 2\int_{T_0}^{T_0+u} \sqrt{|X_s|}\, dW_s\,, \quad u \geq 0\,, \tag{5.39}$$

where T_0 denotes the first time the process $\{X_t\}$ hits 0

$$T_0 = \inf\{t \,|\, X_t = 0\}\,.$$

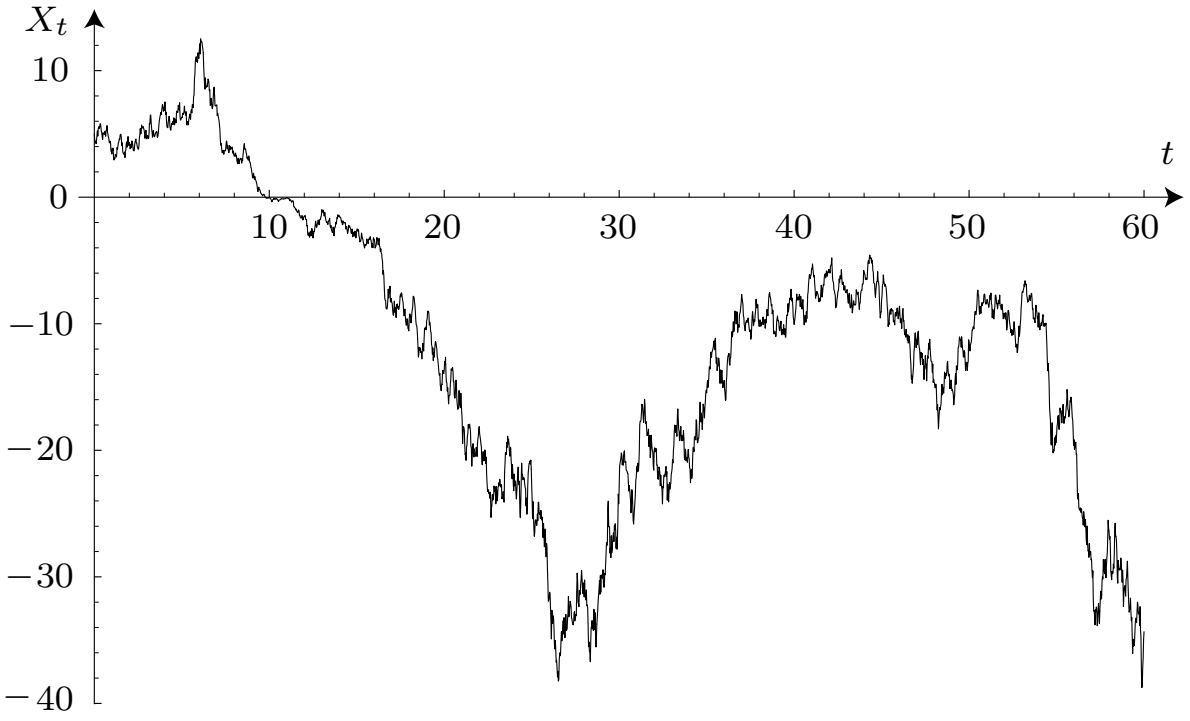

Figure 5.2: *A simulated path of a BES$^{-\gamma}$ process with $\gamma \geq 2$, $x_0 = 5$.*

With the notation $\gamma \equiv -\delta$ we obtain from (5.39)

$$-\tilde{X}_u = \gamma\, u + 2 \int_0^u \sqrt{|\tilde{X}_s|}\, d\tilde{W}_s\,, \quad u \geq 0\,,$$

where $\tilde{W}_s \equiv -(W_{s+T_0} - W_{T_0})$, that is, after the BESQ$_x^{-\gamma}$ process $\{X_t\}$ hits 0, it behaves as a –BESQ$_0^{\gamma}$, $\gamma > 0$. Two simulated paths of a BES$_x^{-\gamma}$ process with $\gamma \geq 2$, respectively $\gamma < 2$, are shown in Figures 5.2 and 5.3.

From the above discussion we deduce that a BESQ$_x^{\delta}$ process with $\delta < 0$ and $x \leq 0$ behaves as a –BESQ$_{-x}^{-\delta}$, especially it never becomes positive. Now we return to the case which motivated us to study Bessel processes with arbitrary δ, $x \in \mathbb{R}$:
For a BESQ$_x^{\delta}$ process with dimension $\delta \geq 0$ and starting point $x \leq 0$, we obtain with the same argument as above, that it behaves as a $-$BESQ$_{-x}^{-\delta}$ process; this means, until it hits 0 for the first time it behaves as a $-$BESQ$_{-x}^{-\delta}$ process, and after that it behaves as a BESQ$_0^{\delta}$.

An important and well-known property of squared Bessel processes with nonnegative dimensions is the additivity property, see Section 5.1, I.3.

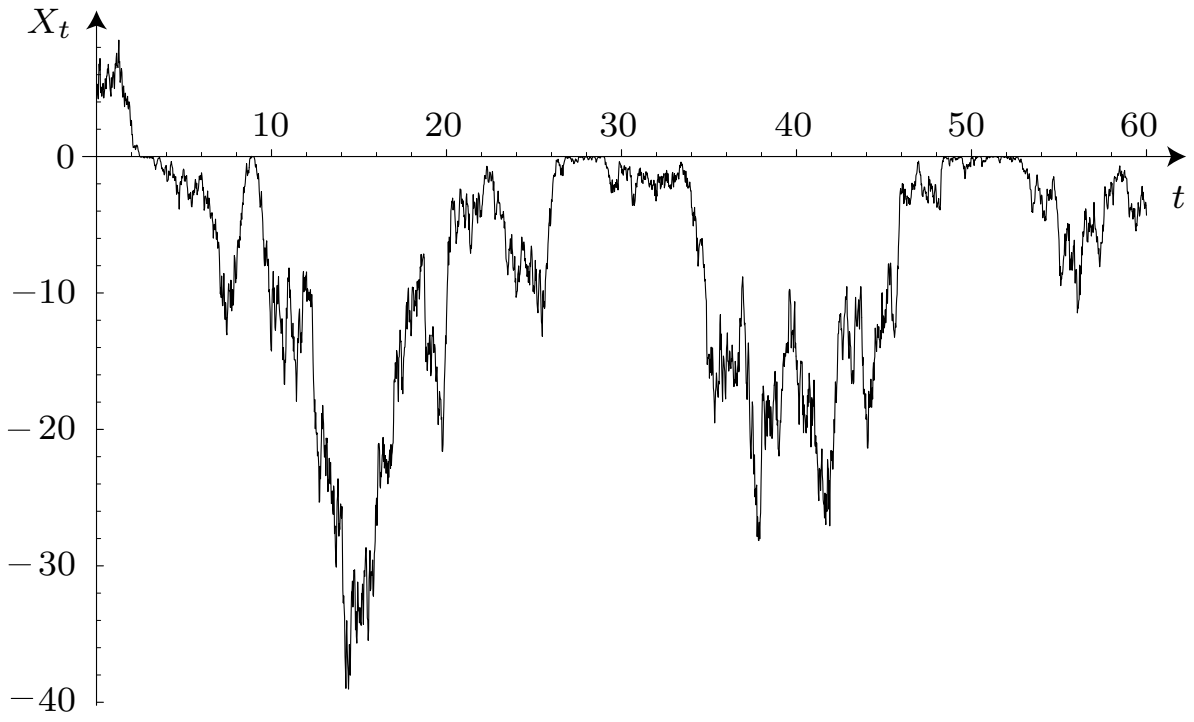

Figure 5.3: *A simulated path of a BES$^{-\gamma}$ process with $\gamma < 2$, $x_0 = 5$.*

We show that the additivity property is no longer true for BESQ$^\delta$ processes with $\delta \in \mathbb{R}$ arbitrary. Consider the BESQ$^\beta_x$ process Z^β and the BESQ$^{\tilde\beta}_y$ process $Z^{\tilde\beta}$, where $\beta > 0$, $\tilde\beta \equiv -\gamma < 0$ with $\beta \geq \gamma$, $x \geq 0$ and $y \leq 0$. Assuming the additivity property holds, would yield:

$$Z^\beta + Z^{\tilde\beta} \overset{\text{(law)}}{=} Z^\beta - Z^\gamma \overset{\text{(law)}}{=} Z^{\beta-\gamma} \geq 0,$$

since $\beta \geq \gamma$. Z^β and Z^γ are independent processes and for the following simple reason $Z^\beta \geq Z^\gamma$ cannot be true: Let X, Y be two independent real-valued random variables with $F_Y(x) < 1$ for all $x \in \mathbb{R}$. Then, if $1 = P(X \geq Y) = E[F_Y(X)]$, we have $F_Y(X) = 1$ a.s. and hence, X has to be infinite, which is a contradiction.

Now we apply the time reversal result Theorem 6 to a BESQ$^{-\gamma}_x$ process $\{X_t\}$ with $\gamma \geq 0$, $x \geq 0$,

$$(X_t,\, t \leq T_0) \overset{\text{(law)}}{=} (\tilde X_{\tilde L_x - t},\, t \leq \tilde L_x), \tag{5.40}$$

where $(\tilde X_t,\, t \geq 0)$ denotes a BESQ$^{4+\gamma}_0$ process and $\tilde L_x = \sup\{t \,|\tilde X_t = x\}$. Our aim is to find the semigroup of a BESQ$^{-\gamma}_x$ process $\{X_t\}$ with $\gamma \geq 0$,

$x \geq 0$. We have to decompose the process (X_t) before and after it hits 0 for the first time separately:

$$E_x^{-\gamma}[f(X_t)] = \underbrace{E_x^{-\gamma}[f(X_t)\,1_{(t<T_0)}]}_{A} + \underbrace{E_x^{-\gamma}[f(X_t)\,1_{(t\geq T_0)}]}_{B}\,.$$

By means of (5.40) we obtain for the first term A

$$A = E_0^{4+\gamma}[f(X_{\tilde{L}_x - t})\,1_{(t<\tilde{L}_x)}]\,,$$

where X is a $\mathrm{BESQ}_0^{4+\gamma}$ process. Since we know (see (5.14))

$$\tilde{L}_x \overset{(\mathrm{law})}{=} \frac{x}{2Z_\nu}\,,$$

where Z_ν is a Gamma variable with parameter $\nu \equiv \frac{4+\gamma}{2} - 1$, see (5.15), we have

$$Q_0^{4+\gamma}(\tilde{L}_x \in dt) = \frac{1}{t\,\Gamma(\nu)} \left(\frac{x}{2t}\right)^{\nu} e^{-\frac{x}{2t}}\,dt \equiv q_x(t)\,dt\,. \tag{5.41}$$

We obtain

$$\begin{aligned} E_0^{4+\gamma}[f(X_{\tilde{L}_x - t})\,1_{(t<\tilde{L}_x)}] &= \int_t^\infty E_0^{4+\gamma}[f(X_{s-t})]\,q_x(s)\,ds \\ &= \int_t^\infty \left(\int_0^\infty f(y)\,p_{s-t}^{4+\gamma}(0,y)\,dy\right) q_x(s)\,ds\,, \end{aligned}$$

where $p_t^\delta(0,y)$, with $\delta > 0$, is the semigroup density $p_t^\delta(x,y)$ in y of BESQ^δ for $x = 0$

$$p_t^\delta(0,y) = \left(\frac{1}{2t}\right)^{\frac{\delta}{2}} \frac{1}{\Gamma(\frac{\delta}{2})}\, y^{\frac{\delta}{2}-1} \exp\left(\frac{-y}{2t}\right)\,.$$

Analogously, for the second term B we have, since $(-X_{t-T_0})$ is a BESQ_0^γ,

$$B = \int_0^t \left(\int_0^\infty f(-y)\,p_{t-s}^\gamma(0,y)\,dy\right) q_x(s)\,ds\,.$$

Putting A and B together we obtain

$$\begin{aligned} E_x^{-\gamma}[f(X_t)] = &\int_t^\infty \left(\int_0^\infty f(y)\,p_{s-t}^{4+\gamma}(0,y)\,dy\right) q_x(s)\,ds \\ &+ \int_0^t \left(\int_0^\infty f(-y)\,p_{t-s}^\gamma(0,y)\,dy\right) q_x(s)\,ds\,, \end{aligned}$$

and hence, the semigroup density $\hat{p}_t^{(-\gamma)}$ in y of a BESQ$^{-\gamma}$ process is equal to

$$\hat{p}_t^{(-\gamma)}(x,y) = \int_t^\infty p_{s-t}^{4+\gamma}(0,y)\, q_x(s)\, ds\,,$$

for $y > 0$, and

$$\hat{p}_t^{(-\gamma)}(x,y) = \int_0^t p_{t-s}^{\gamma}(0,|y|)\, q_x(s)\, ds\,, \tag{5.42}$$

for $y < 0$. We want to find the semigroup density $\hat{p}_t^{(-\gamma)}$ more explicitly. Consider the case $y > 0$:

$$\begin{aligned}\hat{p}_t^{(-\gamma)}(x,y) &= g(x,y,\gamma)\int_t^\infty (s(s-t))^{-(2+\frac{\gamma}{2})} \exp\left[-\frac{1}{2}\left[\frac{y}{s-t}+\frac{x}{s}\right]\right] ds \\ &= g(x,y,\gamma)\int_1^\infty \frac{t^{-3-\gamma}}{(u(u-1))^{(2+\frac{\gamma}{2})}} \exp\left[-\frac{1}{2}\left[\frac{y}{t\,(u-1)}+\frac{x}{tu}\right]\right] du,\end{aligned}$$

where

$$g(x,y,\gamma) = \Gamma^{-2}\left(1+\frac{\gamma}{2}\right)\frac{(xy)^{1+\frac{\gamma}{2}}}{2^{(2+\gamma)}(2+\gamma)}\,.$$

With

$$h(x,y,\gamma,t) \equiv g(x,y,\gamma)\, t^{-3-\gamma},\ m \equiv 2+\frac{\gamma}{2},\ \alpha \equiv \frac{y}{2t},\ \beta \equiv \frac{x}{2t} \tag{5.43}$$

we obtain

$$\begin{aligned}\hat{p}_t^{(-\gamma)}(x,y) &= h(x,y,\gamma,t)\int_1^\infty \frac{du}{(u(u-1))^m} \exp\left(-\left(\frac{\alpha}{u-1}+\frac{\beta}{u}\right)\right) \\ &= h(x,y,\gamma,t)\int_0^1 \frac{s^{2(m-1)}}{(1-s)^m} \exp\left(-\beta s - \frac{\alpha s}{1-s}\right) ds\,,\end{aligned}$$

and hence we have

Proposition 2 *The semigroup density $\hat{p}_t^{(-\gamma)}$ in y of a BESQ$^{-\gamma}$ process is equal to*

$$\hat{p}_t^{(-\gamma)}(x,y) = h(x,y,\gamma,t)\, e^{\alpha-\beta}\int_0^1 \frac{(1-w)^{2(m-1)}}{w^m} \exp\left(\beta w - \frac{\alpha}{w}\right) dw \tag{5.44}$$

for $y > 0$ with (5.43). Analogously, in the case $y < 0$ we have

$$\hat{p}_t^{(-\gamma)}(x,y) = k(x,y,\gamma,t)e^{-\alpha-\beta}\int_0^\infty \frac{(w+1)^{2m}}{w^m}\exp\left(-\beta w - \frac{\alpha}{w}\right)dw\,, \tag{5.45}$$

with

$$k(x,y,\gamma,t) \equiv \Gamma^{-2}\left(\frac{\gamma}{2}\right)\frac{2^{-\gamma}}{\gamma}\,x^{\frac{\gamma}{2}+1}\,|y|^{\frac{\gamma}{2}-1}\,t^{-\gamma-1},$$

and

$$m \equiv \frac{\gamma}{2}\,,\ \alpha \equiv \frac{|y|}{2t}\,,\ \beta \equiv \frac{x}{2t}\,.$$

As an application of the time reversal result (5.40) we make some computations concerning the law $Q_x^{-\gamma}$ of a $\mathrm{BESQ}_x^{-\gamma}$ process $(Z_t)_{t\geq 0}$ with $\gamma, x > 0$. Since $(-Z_{T_0+t}, t \geq 0)$ is a BESQ_0^{γ} process, independent of the past of (Z_t) up to T_0, we have

$$\begin{aligned} Q_x^{-\gamma}&\left[\exp(-\int_0^\infty f(u)\,Z_u\,du)\right] \\ &= \int_0^\infty Q_x^{-\gamma}\left[\exp\,(-\int_0^t f(u)\,Z_u\,du)\middle|\,T_0 = t\right] \\ &\quad\times Q_0^{\gamma}\left[\exp(\int_0^\infty f(t+v)\,X_v\,du)\right]\,Q_x^{-\gamma}[T_0 \in dt]\,, \end{aligned}$$

with a BESQ_0^{γ} process $(X_v)_{v\geq 0}$ and a positive Borel function f. From the time reversal result (5.40) we obtain

$$\begin{aligned} &Q_x^{-\gamma}\left[\exp\,(-\int_0^t f(u)\,Z_u\,du)\,\middle|\,T_0 = t\right] \\ &= Q_0^{4+\gamma}\left[\exp\,(-\int_0^t f(t-u)X_u du)\,\middle|\,L_x = t\right], \end{aligned}$$

and with the equalities

$$\begin{aligned} Q_0^{4+\gamma}&\left[\exp\,(-\int_0^t f(t-u)X_u du)\,\middle|\,L_x = t\right] \\ &= Q_0^{4+\gamma}\left[\exp\,(-\int_0^t f(t-u)X_u du)\,\middle|\,X_t = x\right] \\ &= Q_x^{4+\gamma}\left[\exp\,(-\int_0^t f(t-u)X_{t-u} du)\,\middle|\,X_t = 0\right] \\ &= Q_x^{4+\gamma}\left[\exp\,(-\int_0^t f(u)X_u du)\,\middle|\,X_t = 0\right] \end{aligned}$$

finally, we have with (5.41),

$$Q_x^{-\gamma}\left[\exp(-\int_0^\infty f(u)\,Z_u\,du)\right] = \int_0^\infty Q_x^{4+\gamma}\left[\exp\,(-\int_0^t f(u)X_u du)\,\middle|\,X_t=0\right] \times Q_0^\gamma\left[\exp(\int_0^\infty f(t+v)\,X_v\,dv)\right]\,q_x(t)\,dt\,. \quad (5.46)$$

Our aim is to determine

$$Q_x^{4+\gamma}\left[\exp\,(-\int_0^t f(u)X_u du)\,\middle|\,X_t=0\right]$$

and

$$Q_0^\gamma\left[\exp(\int_0^\infty f(t+v)\,X_v\,dv)\right]$$

more explicitly. We will need

Lemma 2 *Consider a function $h:\mathbb{R}^+\to\mathbb{R}$ in L^1 with $h(x)\le c$ for all $x\in\mathbb{R}^+$ and continuous in a neighbourhood of 0. Then*

$$\lim_{\lambda\to\infty}\int_0^\infty \lambda e^{-\lambda y}h(y)dy = h(0)\,.$$

Proof: Without loss of generality consider $h(0)=0$. For $\varepsilon>0$ choose $\delta>0$ with $h(s)\in(-\varepsilon,\varepsilon)$ for all $s\le\delta$. Then:

$$\left|\int_0^\infty \lambda e^{-\lambda y}h(y)\,dy\right| < \left|\int_0^\delta \lambda e^{-\lambda y}h(y)\,dy\right| + \left|\int_\delta^\infty \lambda e^{-\lambda y}h(y)\,dy\right| < \varepsilon(1-e^{-\lambda\delta}) + ce^{-\lambda\delta} \overset{\lambda\to\infty}{\longrightarrow} \varepsilon\,.$$

□

Applying Lemma 2 we deduce that the ratio

$$\frac{Q_x^\delta(\exp(-\lambda X_t - \int_0^t f(u)X_u\,du))}{Q_x^\delta(\exp(-\lambda X_t))} = \frac{\lambda\int_0^\infty e^{-\lambda y}q_x^\delta(x,y)Q_x^\delta(\exp(-\int_0^t f(u)X_u\,du)|X_t=y)\,dy}{\lambda\int_0^\infty e^{-\lambda y}q_t^\delta(x,y)\,dy}, \quad (5.47)$$

as λ tends to infinity, converges to

$$Q_x^\delta(\exp\{-\int_0^t f(u)X_u\,du\}|X_t=0)\,.$$

Let us consider the numerator and denominator of the ratio (5.47) separately. Via formula (*1.h*) in Pitman–Yor [88] and Theorem (2.1) in Chapter XI in Revuz–Yor [89] we know

Lemma 3 *Consider a positive function $f:\mathbb{R}^+\to\mathbb{R}^+$ with $\int_0^n f(s)ds<\infty$ for all n. Then*

$$Q_x^\delta\left[\exp(-\int_0^t X_u f(u)du-\frac{\lambda}{2}X_t)\right]$$
$$=(\psi_f'(t)+\lambda\psi_f(t))^{-\delta/2}\exp\frac{x}{2}\left[\phi_f'(0)-\frac{(\phi_f'+\lambda\phi_f)(t)}{(\psi_f'+\lambda\psi_f)(t)}\right], \quad (5.48)$$

where ϕ_f, positive and decreasing, is the unique solution of the Sturm-Liouville equation

$$\frac{1}{2}\phi_f''(s)=f(s)\phi_f(s),$$

with $s\in(0,\infty)$, $\phi_f(0)=1$, and where

$$\psi_f(t)\equiv\phi_f(t)\int_0^t\frac{ds}{\phi_f^2(s)}\,.$$

Furthermore

$$Q_x^\delta\left(\exp(-\int_0^\infty X_u f(u)du)\right)=\phi_f(\infty)^{\delta/2}\exp\left(\frac{x}{2}\phi_f'(0)\right), \quad (5.49)$$

where $\phi_f(\infty)$ is the limit at infinity and $\phi_f'(0)$ is the right derivative in 0.

From the formula following Cor.(1.3) in Chapter XI in Revuz–Yor [89] we know

$$Q_x^\delta(\exp(-\frac{\lambda}{2}X_t))=(1+\lambda t)^{-\frac{\delta}{2}}\exp\left(\frac{-\frac{\lambda}{2}x}{1+\lambda t}\right), \quad (5.50)$$

and together with (5.48) we obtain

$$Q_x^\delta(\exp(-\int_0^t f(u)X_u du)|X_t = 0) = \left(\frac{t}{\psi_f(t)}\right)^{\delta/2} \exp\left(\frac{x}{2}(\phi_f'(0) - \frac{\phi_f(t)}{\psi_f(t)} + \frac{1}{t})\right). \quad (5.51)$$

Now we will determine $Q_x^\delta(\exp(\int_0^\infty X_s f(t+s)ds))$ more explicitly. Note that the function

$$\phi_t(s) \equiv \frac{\phi_f(t+s)}{\phi_f(t)}$$

solves

$$\frac{1}{2}\phi_t''(s) = f(t+s)\phi_t(s),$$

and that formula (5.49) leads to

$$Q_x^\delta\left(\exp - \int_0^\infty X_s f(t+s)ds)\right) = \phi_t(\infty)^{\delta/2} \exp\left(\frac{x}{2}\phi_t'(0)\right) = \left(\frac{\phi_f(\infty)}{\phi_f(t)}\right)^{\delta/2} \exp\left(\frac{x}{2}\frac{\phi_f'(t)}{\phi_f(t)}\right).$$

Theorem 4 *For a positive decreasing function $f : \mathbb{R}^+ \to \mathbb{R}^+$, such that*

$$Q_x^\delta\left(\exp\int_0^\infty X_s f(s)ds\right) < \infty,$$

there exists a function $\tilde{\phi}$ which is on $(0,\infty)$ the unique solution of

$$\frac{1}{2}\tilde{\phi}''(s) = -f(s)\tilde{\phi}(s),\ \tilde{\phi}(0) = 1,$$

such that

$$Q_x^\delta\left(\exp\int_0^\infty X_u f(t+u)du\right) = \left(\frac{\tilde{\phi}(\infty)}{\tilde{\phi}(t)}\right)^{\delta/2} \exp\left(\frac{x}{2}\frac{\tilde{\phi}'(t)}{\tilde{\phi}(t)}\right).$$

Proof: We want to determine functions $\alpha(t)$ and $\beta(t)$ with

$$Q_x^\delta\left(\exp\int_0^\infty X_u f(t+u)du\right) = (\beta(t))^{\delta/2} \exp(x\alpha(t)) \quad (5.52)$$

Using the Markov property we have

$$Q_x^\delta\left(\exp\int_0^\infty X_u f(u)du|\mathcal{F}_t\right) = \exp\left(\int_0^t X_u f(u)du\right) Q_{X_t}^\delta\left(\exp\int_0^\infty X_u f(t+u)du\right).$$

Denote

$$h(x,t;\omega) \equiv (\beta(t))^{\delta/2}\exp(x\alpha(t))\exp\left(\int_0^t X_u f(u)du\right).$$

Since $Q_x^\delta\left(\exp\int_0^\infty f(u)X_u du|\mathcal{F}_t\right)$ is a martingale, with Itô's formula we see that the functions α and β have to fulfill the following condition in terms of h

$$\frac{\partial}{\partial t}h(x,t) + 2x\frac{\partial^2}{\partial x^2}h(x,t) + \delta\frac{\partial}{\partial x}h(x,t) = 0.$$

This implies

$$\alpha'(t) + f(t) + 2\alpha^2(t) = 0,\ \alpha(t) + \frac{\beta'(t)}{2\beta(t)} = 0.$$

Hence

$$\alpha(t) = -\frac{\beta'(t)}{2\beta(t)},\ f(t) = \frac{\beta''(t)}{2\beta(t)} - \left(\frac{\beta'(t)}{\beta(t)}\right)^2.$$

Consider the function $\tilde{\phi}:\mathbb{R}^+\to\mathbb{R}^+$ with

$$\tilde{\phi}(s) \equiv \frac{\beta(0)}{\beta(s)}.$$

Then $\tilde{\phi}$ is on $(0,\infty)$ the unique solution of

$$\frac{1}{2}\tilde{\phi}''(t) = -f(t)\tilde{\phi}(t),\ \tilde{\phi}(0) = 1.$$

We obtain

$$\alpha(t) = \frac{\tilde{\phi}'(t)}{2\tilde{\phi}(t)}.$$

From (5.52) we deduce by dominated convergence that $\lim_{t\to\infty}\beta(t) = 1$, and hence $\tilde{\phi}(\infty) \equiv \lim_{t\to\infty}\tilde{\phi}(t) = \beta(0)$. Thus

$$Q_x^\delta\left(\exp\int_0^\infty f(t+u)X_u du\right) = \left(\frac{\tilde{\phi}(\infty)}{\tilde{\phi}(t)}\right)^{\delta/2}\exp\left(\frac{x}{2}\frac{\tilde{\phi}'(t)}{\tilde{\phi}(t)}\right). \qquad \square$$

Altogether we obtain for (5.46) with (5.51) and Theorem 4:

Theorem 5

$$Q_x^{-\gamma}\left(\exp(-\int_0^\infty Z_u f(u)du)\right) = \frac{1}{\Gamma(\frac{\gamma}{2}+1)}\left(\frac{x}{2}\right)^{\frac{\gamma}{2}+1} \times \int_0^\infty \left(\frac{1}{\psi_f(t)}\right)^{\frac{\gamma}{2}+2}\left(\frac{\tilde\phi(\infty)}{\tilde\phi(t)}\right)^{\gamma/2}\exp\left(\frac{x}{2}\left(\phi_f'(0)-\frac{\phi_f(t)}{\psi_f(t)}\right)\right)dt$$

with ϕ_f and ψ_f from Lemma 3, and $\tilde\phi$ from Theorem 4.

As an interesting example consider

$$f(u) \equiv \frac{\lambda^2}{2}\,1_{[0,a]}(u)\,, \quad a,\,\lambda > 0\,.$$

We know (see Pitman–Yor [87], p. 432 (2.m))

$$Q_x^{4+\gamma}\left[\exp\left(-\frac{\lambda^2}{2}\int_0^t X_u du\right)\middle|\, X_t = 0\right] = \left[\frac{\lambda t}{\sinh(\lambda t)}\right]^{\frac{4+\gamma}{2}}\exp\left[-\frac{x}{2t}(\lambda t\coth(\lambda t)-1)\right],$$

and hence we have

$$\begin{aligned} &Q_x^{-\gamma}\left[\exp(-\frac{\lambda^2}{2}\int_0^a Z_t dt)\right] \\ &= \int_0^a \left(\tfrac{\lambda t}{\sinh(\lambda t)}\right)^{\frac{4+\gamma}{2}}\exp\left[-\frac{x}{2t}(\lambda t\coth(\lambda t)-1)\right] \\ &\qquad \times Q_0^\gamma\left[\exp(\frac{\lambda^2}{2}\int_0^{a-t} X_v dv)\right] q_x(t)\,dt \\ &\quad + \int_a^\infty Q_x^{4+\gamma}\left[\exp\left(-\frac{\lambda^2}{2}\int_0^a X_u du\right)\middle|\, X_t = 0\right] q_x(t)\,dt\,. \end{aligned}$$

We investigate more deeply

$$Q^{4+\gamma}_{\substack{x\to 0\\(t)}}\left[\exp(-\frac{\lambda^2}{2}\int_0^a X_u du)\right] \equiv Q_x^{4+\gamma}\left[\exp(-\frac{\lambda^2}{2}\int_0^a X_u du)\middle|\, X_t = 0\right],$$

where a continuous process with law $Q^{4+\gamma}_{\substack{x\to 0\\(t)}}$ is called the *Squared Bessel Bridge* from x to 0 over $[0,t]$. As for (Squared) Bessel Bridges we refer to Revuz–Yor [89], XI.3. For $\mathcal{A}\in\mathcal{F}_a$ we have

$$\begin{aligned} Q^{4+\gamma}_{\substack{x\to 0\\(t)}}(\mathcal{A}) &= \lim_{y\downarrow 0}\frac{E^{4+\gamma}_x\left(1_{\mathcal{A}}1_{[0,y]}(X_t)\right)}{Q^{4+\gamma}_x(X_t\in[0,y])} \\ &= \lim_{y\downarrow 0}\frac{E^{4+\gamma}_x\left[1_{\mathcal{A}}E^{4+\gamma}_x(1_{[0,y]}(X_t)|\mathcal{F}_a)\right]}{Q^{4+\gamma}_x(X_t\in[0,y])} \\ &= \lim_{y\downarrow 0}E^{4+\gamma}_x\left[1_{\mathcal{A}}\frac{Q^{4+\gamma}_{X_a}(X_{t-a}\in[0,y])}{Q^{4+\gamma}_x(X_t\in[0,y])}\right], \end{aligned}$$

and

$$\lim_{y\downarrow 0}\frac{Q^{4+\gamma}_{X_a}(X_{t-a}\in[0,y])}{Q^{4+\gamma}_x(X_t\in[0,y])}=\frac{p^{4+\gamma}_{t-a}(X_a,0)}{p^{4+\gamma}_t(x,0)}\equiv\psi(a,X_a)\,,$$

where $p^\delta_t(x,y)$ is the transition density of BESQ^δ, $\delta>0$. Hence we have

$$Q^{4+\gamma}_{\substack{x\to 0\\(t)}}\left[\exp(-\frac{\lambda^2}{2}\int_0^a X_u du)\right]=Q^{4+\gamma}_x\left[\psi(a,X_a)\exp(-\frac{\lambda^2}{2}\int_0^a X_u du)\right].$$

From the proof of the theorem in 0.5 in Yor [102] we obtain

$$\lim_{y\downarrow 0}\frac{p^{4+\gamma}_{t-a}(X_a,y)}{p^{4+\gamma}_t(0,y)}\cdot\frac{p^{4+\gamma}_t(0,y)}{p^{4+\gamma}_t(x,y)}=\left(\frac{t}{t-a}\right)^{\frac{4+\gamma}{2}}\exp\left[-\frac{X_a}{2(t-a)}\right]\exp\left[\frac{x}{2t}\right].$$

Formula (2.k) in Pitman–Yor [87] gives us

$$\begin{aligned} &Q^{4+\gamma}_x\left[\exp\left[-\frac{X_a}{2(t-a)}-\frac{\lambda^2}{2}\int_0^a X_u\,du\right]\right]= \\ &\left[\cosh(\lambda a)+\frac{1}{\lambda(t-a)}\sinh(\lambda a)\right]^{-\frac{4+\gamma}{2}}\exp\left\{-\frac{\lambda x}{2}\cdot\frac{1+\frac{\coth(\lambda a)}{\lambda(t-a)}}{\coth(\lambda a)+\frac{1}{\lambda(t-a)}}\right\}. \end{aligned}$$

Hence

$$\begin{aligned} Q^{4+\gamma}_{\substack{x\to 0\\(t)}}\left[\exp(-\frac{\lambda^2}{2}\int_0^a X_u du)\right]&=\left[\left(\frac{t-a}{t}\right)\cosh(\lambda a)+\frac{1}{\lambda t}\sinh(\lambda a)\right]^{-\frac{4+\gamma}{2}} \\ &\quad\cdot\exp\left\{\frac{x}{2t}-\frac{\lambda x}{2}\frac{1+\frac{\coth(\lambda a)}{\lambda(t-a)}}{\coth(\lambda a)+\frac{1}{\lambda(t-a)}}\right\}. \end{aligned}$$

If in addition $\lambda(a-t) < \frac{\pi}{2}$,

$$Q_0^\gamma \left[\exp(\frac{\lambda^2}{2} \int_0^{a-t} X_v dv) \right] = \cos(\lambda(a-t))^{-\frac{\gamma}{2}},$$

and finally, we have (with $\nu = \frac{\gamma}{2} + 1$)

$$\begin{aligned} & Q_x^{-\gamma} \left[\exp(-\frac{\lambda^2}{2} \int_0^a Z_t dt) \right] \\ & = \left(\frac{x}{2}\right)^\nu \frac{1}{\Gamma(\nu)} \left(\lambda^{\nu+1} \int_0^a \exp\left[-\frac{x}{2} \lambda \coth(\lambda t) \right] \frac{\cos(\lambda(a-t))^{-\nu+1}}{\sinh(\lambda t)^{\nu+1}} \, dt \right. \\ & \quad + \int_a^\infty [(t-a)\cosh(\lambda a) + \frac{1}{\lambda}\sinh(\lambda a)]^{-\nu-1} \\ & \qquad \left. \exp\left[-\frac{\lambda x}{2} \frac{1 + \frac{\coth(\lambda a)}{\lambda(t-a)}}{\coth(\lambda a) + \frac{1}{\lambda(t-a)}} \right] dt \right). \end{aligned}$$

Extension

Now we extend the results for squares of Bessel processes with negative dimensions to a wider class of processes. We generalize the definition of a squared δ-dimensional radial Ornstein-Uhlenbeck process with parameter $-\lambda$ and starting point z in (5.9), by considering arbitrary δ, $z \in \mathbb{R}$.

Definition 4 *The solution to the stochastic differential equation*

$$dX_t = (a + b\, X_t)\, dt + 2\sqrt{|X_t|}\, dW_t, \quad X_0 = x, \tag{5.53}$$

where a, b, $x \in \mathbb{R}$ and $\{W_t\}$ is a one-dimensional Brownian motion, is called a squared radial Ornstein–Uhlenbeck process.

In the following, we consider the case $a < 0$ and $x > 0$ which corresponds for $b = 0$ to a BESQ_x^a process with negative dimension a. We call ${}^bQ_x^a$ the law on $C(\mathbb{R}_+, \mathbb{R})$ and for $b = 0$, as before, simply Q_x^a. Via Girsanov transformation we obtain the relationship

Lemma 4

$${}^bQ_x^a\big|_{\mathcal{F}_t} = \exp\left[\frac{b}{2} \int_0^t \sigma(X_s)\, dW_s - \frac{b^2}{8} \int_0^t |X_s|\, ds \right] Q_x^a\big|_{\mathcal{F}_t}, \tag{5.54}$$

where $\sigma(x) = \sqrt{|x|}\,\mathrm{sgn}\,(x)$.

Proof: Consider the Q_x^a-process

$$dX_t = a\,dt + 2\sqrt{|X_t|}\,dW_t\,, \quad X_0 = x\,.$$

Via Girsanov's theorem $\tilde{W}_t = W_t - \frac{b}{2}\int_0^t \operatorname{sgn}(X_s)\sqrt{|X_s|}\,ds$ is a Brownian motion under the absolutely continuous probability with density

$$D_t \equiv \exp\left[\frac{b}{2}\int_0^t \operatorname{sgn}(X_s)\sqrt{|X_s|}\,dW_s - \frac{b^2}{8}\int_0^t |X_s|\,ds\right],$$

with respect to Q_x^a, from which the result follows. □

We may also write the stochastic integral $\int_0^t \operatorname{sgn}(X_s)\sqrt{|X_s|}\,dW_s$ in a simpler form, since we have from Itô's formula:

$$|X_t| = |x| + \int_0^t \operatorname{sgn}(X_s)\,(a\,ds + 2\sqrt{|X_s|}\,dW_s) + L_t^0(X)\,,$$

where $L_t^0(X)$ is the semimartingale local time of X in 0. For $L_t^0(X)$ we obtain:

$$\begin{aligned} L_t^0(X) &= \lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}\int_0^t 1_{[0,\varepsilon]}(X_s)\,d\langle X,X\rangle_s \qquad ([89],\text{ p. 212, Cor. 1.9}) \\ &= \lim_{\varepsilon\downarrow 0}\frac{4}{\varepsilon}\int_0^t |X_s|\,1_{[0,\varepsilon]}(X_s)\,ds \\ &\le \lim_{\varepsilon\downarrow 0}\left(4\int_0^t 1_{[0,\varepsilon]}(X_s)\,ds\right) = 0\,. \end{aligned}$$

Hence we have

$$\int_0^t \operatorname{sgn}(X_s)\sqrt{|X_s|}\,dW_s = \frac{1}{2}\left[\,|X_t| - |x| - a\int_0^t \operatorname{sgn}(X_s)\,ds\right].$$

Thus, (5.54) takes the form

$$\left.\frac{{}^bQ_x^a}{Q_x^a}\right|_{\mathcal{F}_t} \equiv \exp\left[\frac{b}{4}\left(|X_t| - |x| - a\int_0^t \operatorname{sgn}(X_s)\,ds\right) - \frac{b^2}{8}\int_0^t |X_s|\,ds\right]. \tag{5.55}$$

Now we can apply formula (5.55) to obtain the conditional expectation formula

$$\begin{aligned} {}^bq_t^a(x,y) &= q_t^a(x,y)\,\exp\left(\frac{b}{4}\,(|y| - |x|)\right) \\ &\times Q_x^a\left[\exp\left(-\frac{ab}{4}\int_0^t \operatorname{sgn}(X_s)\,ds - \frac{b^2}{8}\int_0^t |X_s|\,ds\right)\,\middle|\, X_t = y\right]. \end{aligned} \tag{5.56}$$

The case $y > 0$ corresponds to $t < T_0$, and $y \leq 0$ to $t \geq T_0$. For $y > 0$, (5.56) reduces to

$$
\begin{aligned}
{}^b q_t^a(x,y) = q_t^a(x,y) \exp\left(\frac{b}{4}(y-x) - \frac{abt}{4}\right) \\
\times Q_x^a\left[\exp\left(-\frac{b^2}{8}\int_0^t X_s\,ds\right)\,\middle|\, X_t = y\right]. \qquad (5.57)
\end{aligned}
$$

Using the time-space transformation from a BESQ^a process (X_t^a) to a squared radial Ornstein–Uhlenbeck process $({}^bX_t^a)$

$$
{}^bX_t^a = e^{bt}\, X^a_{\left(\frac{1-e^{-bt}}{b}\right)},
$$

we also have together with the relationship (5.56) the following

$$
{}^b q_t^a(x,y) = e^{-bt} q^a_{\left(\frac{1-e^{-bt}}{b}\right)}(x,\, e^{-bt}y), \qquad (5.58)
$$

from which ${}^b q_t^a(x,y)$ is obtained since we know $q_t^a(x,y)$, the semi-group density in y of a BESQ^a process, $a < 0$, given by (5.44) for $y > 0$ and given by (5.45) for $y < 0$. Hence we obtain for $y > 0$ from (5.57)

$$
\begin{aligned}
&Q_x^a\left[\exp\left(-\frac{b^2}{8}\int_0^t X_s\,ds\right)\,\middle|\, X_t = y\right] \\
&= q^a_{\left(\frac{1-e^{-bt}}{b}\right)}(x, e^{-bt}y)\exp\left(-bt - \frac{b}{4}(y-x) + \frac{abt}{4}\right)/q_t^a(x,y),
\end{aligned}
$$

with $q_t^a(x,y)$ given by (5.44). For $y < 0$ we have with (5.56)

$$
\begin{aligned}
&Q_x^a\left[\exp\left(-\frac{ab}{4}\int_0^t \operatorname{sgn}(X_s)\,ds - \frac{b^2}{8}\int_0^t |X_s|\,ds\right)\,\middle|\, X_t = y\right] \\
&= q^a_{\left(\frac{1-e^{-bt}}{b}\right)}(x,\, e^{-bt}y)\exp\left(-bt - \frac{b}{4}(|y| - |x|)\right)/q_t^a(x,y),
\end{aligned}
$$

with $q_t^a(x,y)$ given by (5.45).

5.4 Inhomogeneous Bessel Processes

In mathematical finance stock prices are usually assumed to follow the exponential of Brownian motion with drift $\exp(B_t + \nu t)$. Bessel processes

appear in this context, since the exponential of Brownian motion with homogeneous drift is a time-changed Bessel process. Indeed, we know, see Section 5.1, I.4 (5.5),

$$\exp(B_t + \nu t) = X^{(\nu)} \left(\int_0^t \exp 2(B_s + \nu s)\, ds \right), \quad t \geq 0\,,$$

(Lamperti [64], Williams [99]), where $X^{(\nu)}$ is a Bessel process with index ν. Note that using the Brownian scaling property immediately also leads to a representation of the process $\exp(aB_t + \nu t)$, $a > 0$, in terms of Bessel processes. Geman and Yor [35] investigate problems including integrals of geometric Brownian motion, and by using (5.5), i.e. by using properties of Bessel processes, they find for instance a formula for the LT of an Asian option price.

As an extension now we investigate the exponential of Brownian motion with inhomogeneous drift $\exp(Z_t)$, where

$$Z_t = B_t + f(t)$$

with

$$f(t) = \int_0^t g(s) ds,$$

$t \geq 0$, where $g \in L^1$. By Itô's formula we obtain

$$\exp(Z_t) = 1 + \int_0^t \exp(Z_s) \left(\frac{1}{2} + g(s) \right) ds + \int_0^t \exp(Z_s)\, dB_s\,.$$

Consider the process $R_u \equiv \exp(Z_{H_u})$ with the time change

$$H_u \equiv \int_0^u \frac{ds}{R_s^2} = \int_0^u \frac{ds}{\exp(2Z_{H_s})}\,.$$

Note that with

$$A_s \equiv \int_0^s \exp(2Z_u) du$$

we have $A_{H_t} = t$. With $\beta_{(A_t)} \equiv \int_0^t \exp(Z_s)\, dB_s$ for a standard Brownian motion β we obtain

$$\begin{aligned} R_u = \exp(Z_{H_u}) &= 1 + \beta_u + \int_0^{H_u} \exp(Z_s) \left(\frac{1}{2} + g(s) \right) ds \\ &= 1 + \beta_u + \int_0^u \exp(Z_{H_v}) \left(\frac{1}{2} + g(H_v) \right) dH_v\,, \end{aligned}$$

and since $dH_v = \frac{dv}{\exp(2Z_{H_v})}$, we have

$$R_u = 1 + \beta_u + \int_0^u \frac{1}{R_v}\left(\frac{1}{2} + g(H_v)\right) dv\,. \tag{5.59}$$

The pair $(R_t, H_t; t \geq 0)$ is a Markov process with infinitesimal generator

$$\frac{1}{2}\frac{\partial^2}{\partial r^2} + \frac{1}{r}\left(\frac{1}{2} + g(h)\right)\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial}{\partial h}\,.$$

From (5.59) we obtain via Girsanov's theorem immediately the relationship

$$P_r^g\big|_{\mathcal{R}_t} = \exp\left\{\int_0^t \frac{g(H_u)}{R_u} d\beta_u - \frac{1}{2}\int_0^t \left(\frac{g(H_u)}{R_u}\right)^2 du\right\} \cdot P_r^0\big|_{\mathcal{R}_t}\,,$$

where P^g denotes the law of R satisfying (5.59) and P^0 is the law of a two-dimensional Bessel process, and we call such a process R with law P^g an *inhomogeneous* two-dimensional Bessel process.

5.5 Time Reversal

In the previous sections of this chapter we used results on time reversed diffusions as an essential tool, e.g. in the first section as for the distribution of first hitting times by Bessel processes, or in Section 5.3 as for the transition densities of BESQ processes with negative dimensions. In this section we investigate these time reversal results more deeply. As for time reversal we refer to Nagasawa [72, 73], Revuz–Yor [89], Chapter VII, and Dellacherie–Meyer [23]. First, some results on time reversed diffusions in general are reviewed. In Section 5.5.1 we study Doob's h-transform and consider some applications. Using these results, in Section 5.5.2 we check a time reversal theorem by Elworthy–Li–Yor [26], and based on this theorem, we introduce a three-parameters-family of processes in Section 5.5.3.

Consider a transient diffusion (X_t), living on $\mathbb{R}_+$, with $X_0 = x_0 \geq 0$. Denote its last exit time of $a \geq 0$ by $L_a = \sup\{u \,|\, X_u = a\}$, where $\sup \emptyset = 0$. For a fixed, L_a is finite a.s., and for $x_0 < a$, $L_a > 0$ a.s. We

want to identify the time reversed process $(\tilde{X}_t)$, where

$$\tilde{X}_t(\omega) \equiv \begin{cases} X_{L_a(\omega)-t}(\omega)\,, & \text{if } 0 < t < L_a(\omega)\,, \\ \partial & \text{if } L_a(\omega) \le t \text{ or } L_a(\omega) = \infty\,, \end{cases} \tag{5.60}$$

where ∂ denotes the "cemetery", and $\tilde{X}_0(\omega) = X_{L_a(\omega)}(\omega)$, if $0 < L_a(\omega) < \infty$, else $\tilde{X}_0(\omega) = \partial$. We remark that a diffusion cannot only be time reversed at a last exit time, but more generally at a cooptional time, see Nagasawa [72, 73], Revuz–Yor [89], Chapter VII.4. For our purposes it is reasonable to restrict ourselves to last exit times.

We will need the following (see also Itô–McKean [49], p. 149)

Lemma 5 *Let (p_t) be the transition density of a diffusion (X_t) with respect to the speed measure m. Then we have for all x, y the symmetry*

$$p_t(x,y) = p_t(y,x)\,.$$

Proof: Let A denote the infinitesimal generator of the diffusion and let the functions ψ_α and ϕ_α be the fundamental solutions of the generalized differential equation

$$Au = \alpha u, \quad \alpha > 0\,.$$

Call f^+ the right derivative of a function f with respect to the scale function s of the diffusion,

$$f^+(x) = \lim_{h \downarrow 0} \frac{f(x+h) - f(x)}{s(x+h) - s(x)}\,.$$

Then the so called Wronskian w_α is defined as

$$w_\alpha = \psi_\alpha^+(x)\phi_\alpha(x) - \psi_\alpha(x)\phi_\alpha^+(x)$$

and is independent of x. The function

$$g_\alpha(x,y) \equiv \int_0^\infty e^{-\alpha t} p_t(x,y)\, dt \tag{5.61}$$

is called the Green function and

$$g_\alpha(x,y) = \begin{cases} w_\alpha^{-1}\psi_\alpha(x)\phi_\alpha(y), & x \le y, \\ w_\alpha^{-1}\psi_\alpha(y)\phi_\alpha(x), & x \ge y. \end{cases} \tag{5.62}$$

We obtain $g_\alpha(x,y) = g_\alpha(y,x)$, and hence, from the injectivity of the Laplace transform the result follows. □

Furthermore, some assumptions and notations are needed. Call (P_t) the semigroup of (X_t). The time reversed process $(\tilde{X}_t)$ is a Markov process (see Nagasawa [72] and [73], III.3.2, p. 66) and we call its semigroup $(\tilde{P}_t)$. The resolvent G_α of order α, $\alpha > 0$, of the semigroup (P_t), and its analog $\tilde{G}_\alpha$ for $(\tilde{P}_t)$, is

$$G_\alpha(f)(x) = \int_0^\infty e^{-\alpha t}\, P_t f(x)\, dt\,,$$

for all positive Borel functions f. The semigroup (P_t) has a transition density (p_t) with respect to the speed measure m: $P_t(x,dy) = p_t(x,y)\, m(dy)$ (see Itô–McKean [49], p. 149). We have

$$\begin{aligned} G_\alpha(f)(x) = \int_0^\infty e^{-\alpha t} P_t f(x)\, dt &= \int_0^\infty e^{-\alpha t} \left(\int p_t(x,y) f(y)\, m(dy) \right) dt \\ &= \int g_\alpha(x,y) f(y)\, m(dy), \end{aligned}$$

for all positive Borel functions f, with g_α in (5.61). G_0 is the potential kernel of X, and we assume that there is a probability measure μ such that the potential $\nu = \mu G_0$ is a Radon measure, and additionally, that the resolvents $\tilde{G}_\alpha$ and G_α are in duality with respect to ν, that is

$$\langle G_\alpha f, g \rangle_\nu = \langle f, \tilde{G}_\alpha g \rangle_\nu\,,$$

for all $\alpha > 0$ and for all positive Borel functions f and g. Then we have the duality between the semigroups (P_t) and $(\tilde{P}_t)$:

$$\langle P_t f, g \rangle_\nu = \langle f, \tilde{P}_t g \rangle_\nu \tag{5.63}$$

for all positive Borel functions f and g, see Nagasawa [73], §3.2, Revuz–Yor [89], Theorem 4.5, p. 292.

From

$$\mu G_0(A) = \int G_0(x,A) \mu(dx) = \int_A \left(\int g_0(x,y) \mu(dx) \right) m(dy),$$

for every Borel set A, we obtain $\nu = \mu G_0 \ll m$. With the notation

$$h(y) \equiv \frac{d(\mu G_0)}{dm}\,(y) = \int g_0(x,y)\, \mu(dx) \tag{5.64}$$

for m-almost all y we have

$$\begin{aligned}\langle P_t f, g\rangle_\nu &= \int \left(\int f(y)\, p_t(x,y)\, m(dy) \right) g(x)h(x)m(dx) \\ &= \int \left(\frac{1}{h(y)} \int g(x)h(x)\, p_t(x,y) m(dx) \right) f(y)h(y)\, m(dy) \\ &= \int \frac{1}{h(y)} P_t(gh)(y)\, f(y)h(y)\, m(dy)\,,\end{aligned}$$

by Lemma 5, and from duality (5.63) we deduce

$$\tilde{P}_t g(x) = \frac{1}{h(x)} P_t(gh)(x) \tag{5.65}$$

for m-almost all x.

We state a time reversal result proven with the results above (see Sharpe [93], Getoor–Sharpe [37]), which was a useful tool in the previous sections.

Theorem 6 *For X a transient diffusion, living on $\mathbb{R}_+$, starting at 0, $\tilde{X}$ the time reversed process (5.60), starting at a, and $T_0 \equiv \inf\{u | \tilde{X}_u = 0\}$ we have*

$$\{X_u,\, u \le L_a\} \overset{(\text{law})}{=} \{\tilde{X}_{T_0 - u},\, u \le T_0\}\,. \tag{5.66}$$

This implies that the law of L_a for the process X starting at 0 is identical to the law of T_0 for the process $\tilde{X}$ starting at a.

We want to derive a more explicit formula for h in (5.64). We know, see (5.31):

$$P_{x_0}(L_a \in dt) = -\frac{1}{2s(a)}\, p_t(x_0, a)\, dt\,, \tag{5.67}$$

where $X_0 = x_0$, and s is the scale function with $\lim_{y \downarrow 0} s(y) = -\infty$ and $s(\infty) = 0$. Consider the two cases $x_0 \le a$ and $x_0 > a$. If $x_0 \le a$, then

$$-\frac{1}{2\, s(a)} \int_0^\infty p_t(x_0, a)\, dt = 1\,,$$

hence $-2\, s(a) = g_0(x_0, a)$. If $x_0 > a$, then the law of L_a has mass at 0

$$P_{x_0}(L_a = 0) = P_{x_0}(T_a = \infty) = 1 - \frac{s(x_0)}{s(a)}\,,$$

hence,

$$P_{x_0}(0 < L_a) = -\frac{1}{2s(a)} \int_0^\infty p_t(x_0, a)\, dt = \frac{s(x_0)}{s(a)}\,,$$

and $-2\, s(x_0) = g_0(x_0, a)$. We conclude

$$g_0(x_0, a) = -2\, s(x_0 \vee a) \tag{5.68}$$

for $x_0, a \in \mathbb{R}^+$ arbitrary.

Now, assume $x_0 = 0$, that is, the transient process (X_t) starts at 0 and μ is the Dirac measure at 0. Then from (5.64) and (5.68) we obtain

$$h(a) = g_0(0, a) = -2\, s(a). \tag{5.69}$$

As we will see in the next section, the process $(\tilde{X}_t)$ is Doob's h-transform of (X_t).

5.5.1 Doob's h-Transform

Consider a one-dimensional diffusion X, with sample space $(I^{\partial,\infty}, \mathcal{F}^c_\infty)$, where $I \subseteq [-\infty, \infty]$, ∂ denotes the "cemetery" and $I^{\partial,\infty} := \{\omega : [0,\infty) \mapsto I \cup \{\partial\}\}$, $\mathcal{F}^c_\infty := \sigma\{\omega(t) |\, t \geq 0\}$.

Definition 5 *A non-negative measurable function $h : I \mapsto \mathbb{R} \cup \{\infty\}$ is called α-excessive for X, $\alpha \geq 0$, if*

a) $e^{-\alpha t} E_x(h(X_t)) \leq h(x)$, for all $x \in I$, $t \geq 0$,

b) $e^{-\alpha t} E_x(h(X_t)) \to h(x)$, for all $x \in I$ as $t \downarrow 0$.

A 0-excessive function is simply called excessive.

Let h be an α-excessive function for a diffusion X. The life time of a path $\omega \in I^{\partial,\infty}$ is defined by $\zeta(\omega) := \inf\{t \,|\, \omega_t = \partial\}$. We construct a new probability measure P^h by

$$P^h_x\big|_{\mathcal{F}_t} = e^{-\alpha t}\, \frac{h(\omega(t))}{h(x)}\; P_x|_{\mathcal{F}_t}\,, \tag{5.70}$$

for $t < \zeta$ and $x \in I$. The process under the new measure P^h is a regular diffusion and is called *Doob's h-transform* of X. Examples of α-excessive functions are ψ_α and ϕ_α, see Lemma 5, and $g_\alpha(x,\cdot)$, see (5.61). We remark, that all α-excessive functions h (except for $h \equiv 0$ or $h \equiv \infty$) can be expressed in terms of a linear combination of these three functions. As for Doob's h-transform we certainly refer to Doob [24], but also to Dellacherie–Meyer [23]; in presenting Doob's h-transform we followed Borodin–Salminen [10].

As an application we consider Doob's h-transform (5.70) with the excessive function h in (5.69) for the transient diffusion X. As seen in (5.70) the factor -2 cancels, so we simply consider $h \equiv s$, with s the scale function of X. Let us denote the probability measure associated to X by $P^{(\uparrow)}$. We obtain:
Doob's h-transform of the transient diffusion X with $h \equiv s$ the scale function of X, that is, the process under the new measure P^h, is a process which reaches 0 almost surely. For clarity we write $P^{(0)} \equiv P^h$. With this notation we have

$$P_x^{(\uparrow)}\Big|_{\mathcal{F}_t} = \frac{\left(\frac{1}{s}\right)(X_{t\wedge T_0})}{\left(\frac{1}{s}\right)(x)}\; P_x^{(0)}\Big|_{\mathcal{F}_t}\,. \tag{5.71}$$

Note that we obtain

$$\frac{s\,(X_t)}{s(x)}\; P_x^{(\uparrow)}\Big|_{\mathcal{F}_t} \equiv 1_{(t<T_0)}\; P_x^{(0)}\Big|_{\mathcal{F}_t}\,,$$

and hence,

$$Q_t^{(0)} f(x) \equiv \frac{1}{s(x)}\, E_x\left[(fs)(X_t)\right] \equiv \frac{1}{s(x)} P_t^{(\uparrow)}(fs)(x)\,, \tag{5.72}$$

is the semigroup of the process under $P^{(0)}$ killed when it reaches 0. In other words, we have the following time reversal result: Doob's h-transform of the transient process under $P^{(\uparrow)}$ with $h = s$ is the process under $P^{(0)}$ killed at T_0; this is the process $\tilde{X}$ in our former notation.

From formula (5.71) we can obtain explicit formulae of the diffusion processes via Girsanov's theorem. Assume, a process (X_t) under $P_x^{(0)}$ has the form

$$X_t = x + \int_0^t b(X_s)\,ds + \int_0^t a(X_s)\,dB_s, \quad t \le T_0\,.$$

Via Girsanov's theorem we have with (5.71)

$$B_t = \int_0^t \frac{\left(\frac{1}{s}\right)'(X_u)}{\left(\frac{1}{s}\right)(X_u)}\,du + \hat{B}_t = \int_0^t -\left(\frac{s'}{s}\right)(X_u)\,du + \hat{B}_t\,,$$

where $(\hat{B}_t)$ is Brownian motion under $P_x^{(\uparrow)}$. Hence, the process under $P_x^{(\uparrow)}$ has the form

$$X_t = x + \int_0^t \left(b - \left(a\,\frac{s'}{s}\right)\right)(X_u)\,du + \int_0^t a(X_u)\,d\hat{B}_u\,.$$

Note, that if we assume, a process (X_t) under $P_x^{(\uparrow)}$ has the form

$$X_t = x + \int_0^t \beta(X_u)\,du + \int_0^t \alpha(X_u) d\hat{B}_u\,,$$

then analogously we obtain that the process under $P_x^{(0)}$ has the form

$$X_t = x + \int_0^t \left(\beta + \left(\alpha\,\frac{s'}{s}\right)\right)(X_u)\,du + \int_0^t \alpha(X_u)\,dB_u, \quad t \le T_0\,.$$

Let $A^{(0)}$ respectively $A^{(\uparrow)}$ denote the infinitesimal generator of the process under $P^{(0)}$ respectively $P^{(\uparrow)}$. If we want to avoid a Girsanov argument, we can compute the infinitesimal generator of the time-reversed process from (5.72)

$$A^{(0)}(f) = \frac{1}{h} A^{(\uparrow)}(hf)\,,$$

with h the scale function of $X^{(\uparrow)}$. The term $A(hf)$ can be written as

$$A(hf) = hA(f) + fA(h) + \Gamma(h, f)\,,$$

where Γ is the 'Opérateur Carré du Champ', see Kunita [60] and Revuz–Yor [89], p. 326. For instance, if the infinitesimal generator takes the form $A(f) = \frac{1}{2}f'' + bf'$, then $\Gamma(f, g) = f'g'$ for $f, g \in C^2$.

5.5.2 Checking a Time Reversal Theorem in Elworthy–Li–Yor

Considering properties of radial Ornstein–Uhlenbeck processes is one of our main interests: in Section 5.2 we investigated first hitting times of

radial Ornstein–Uhlenbeck processes and in Section 5.3 we also considered negative dimensional squared radial Ornstein–Uhlenbeck processes. Therefore, as an application of the previously stated results, now we want to check a time reversal result for radial Ornstein–Uhlenbeck processes, Theorem 3.7 in Elworthy–Li–Yor [26]. In Section 5.5.3 we will introduce a three-parameters-family of processes based on this theorem. Let ${}^{-\lambda}P_x^\delta$ denote the law of a δ-dimensional radial Ornstein–Uhlenbeck process with parameter $-\lambda$, starting in x. We have the following time reversal result:

Theorem 7 (Elworthy–Li–Yor [26]) *Let $\lambda \geq 0$ and $\delta > 2$. For every bounded measurable function F on path space,*

$$ {}^{\lambda}E_0^\delta[F(Y_{L_a - t},\, t \leq L_a)] = \frac{1}{g(a)}\, {}^{-\lambda}E_a^{(4-\delta)}[F(Y_t,\, t \leq T_0)\, e^{-2\lambda T_0}]\,, \quad (5.73) $$

where $g(a) = {}^{-\lambda}E_a^{(4-\delta)}(e^{-2\lambda T_0})$.

Let (R_t) be a radial Ornstein–Uhlenbeck process with law ${}^{\lambda}P_0^\delta$ and denote by σ its scale function and by A its infinitesimal generator. We know that the infinitesimal generator of the time-reversed process is

$$ \begin{aligned} \frac{1}{\sigma} A(\sigma f) &= \frac{1}{\sigma}\left[\sigma A(f) + f A(\sigma) + \Gamma(\sigma, f)\right] \\ &= A(f) + \frac{1}{\sigma}\Gamma(\sigma, f) = A(f) + \frac{\sigma'}{\sigma} f'\,. \end{aligned} \quad (5.74) $$

Let us check whether this is coherent with Theorem 7. From

$$ \begin{aligned} {}^{\lambda}E_0^\delta[F(Y_{L_a - t},\, t \leq L_a)] &= \frac{1}{g(a)}\, {}^{-\lambda}E_a^{(4-\delta)}[F(Y_t,\, t \leq T_0)\, e^{-2\lambda T_0}] \\ &\equiv \hat{E}_a[F(\hat{Y}_t,\, t \leq \hat{T}_0)] \end{aligned} $$

we obtain

$$ \hat{P}_a\Big|_{\mathcal{F}_t} = D_t \cdot\ {}^{-\lambda}P_a^{(4-\delta)}\Big|_{\mathcal{F}_t}\,, \quad (5.75) $$

with

$$ D_t \equiv \frac{g(Y_{t\wedge T_0})}{g(a)}\, e^{-2\lambda(t\wedge T_0)}\,. $$

Writing the process $(Y_t,\, t \le T_0)$ simply as

$$Y_t = \int_0^t b(Y_s)ds + B_t\,,$$

where B is Brownian motion under ${}^{-\lambda}P_a^{(4-\delta)}$, then we obtain by Girsanov's theorem

$$Y_t = \int_0^t \left(b + \frac{g'}{g}\right)(Y_s)\,ds + \hat{B}_t\,,$$

where $\hat{B}$ is Brownian motion under $\hat{P}_a$. Let A^* denote the infinitesimal generator of the process (Y_t) under ${}^{-\lambda}P_a^{(4-\delta)}$. We have for the new infinitesimal generator $\hat{A}$ of the process $(\hat{Y}_t)$ under $\hat{P}_a$:

$$\hat{A}(f) = A^*(f) + \frac{g'}{g}f'\,.$$

Now, all we have to show is (see (5.74))

$$A(f) + \frac{\sigma'}{\sigma}f' = A^*(f) + \frac{g'}{g}f'\,,$$

that is, we have to identify the drift coefficients:

$$\frac{\delta-1}{2r} + \lambda r + \frac{\sigma'}{\sigma} = \frac{3-\delta}{2r} - \lambda r + \frac{g'}{g}\,. \tag{5.76}$$

The problem is that we do not know g explicitly. But since D_t is a martingale, the following two lemmas will help us.

Lemma 6 *$g(Y_t)\,e^{-2\lambda t}$ is a ${}^{-\lambda}P_a^{(4-\delta)}$ martingale if and only if $g \in D(A)$ and $A(g) = 2\lambda g$.*

Proof: If $g(Y_t)e^{-2\lambda t}$ is a martingale, then we have $P_h g(Y_t)e^{-2\lambda(t+h)} - g(Y_t)e^{-2\lambda t} = 0$ for all $h > 0$, hence $e^{-2\lambda t}P_t g(x) = g(x)$ with $Y_0 = x$ and we have $\frac{d}{dt}P_t g(x)|_{t=0} = 2\lambda g(x)$, that is $g \in D(A)$.

For $g \in D(A)$, the process

$$M_t^g = g(Y_t) - g(Y_0) - \int_0^t Ag(Y_s)ds$$

is a martingale, see Revuz–Yor [89] Proposition 7.1.6. By Itô's lemma we obtain

$$g(Y_t)e^{-2\lambda t} = \int_0^t e^{-2\lambda s} d(g(Y_s)) - 2\lambda \int_0^t e^{-2\lambda s} g(Y_s) ds \,.$$

With the notation: $U \sim V$ if and only if $U_t - V_t$ is a martingale, we obtain:

$$g(Y_t)e^{-2\lambda t} \sim \int_0^t e^{-2\lambda s} (Ag)(Y_s) ds - 2\lambda \int_0^t e^{-2\lambda s} g(Y_s) ds \,,$$

and the claim follows. □

Lemma 7 *$(\frac{g}{\sigma})(Y_t)e^{-2\lambda t}$ is a martingale under ${}^{-\lambda}P_a^{(4-\delta)}$.*

Proof: Consider Doob's h-transform of the process under ${}^{\lambda}P_0^{\delta}$ with $h = \sigma$ its scale function, or put differently (see (5.71))

$${}^{\lambda}P_0^{\delta}\big|_{\mathcal{F}_t} = M_t \cdot \hat{P}_a\Big|_{\mathcal{F}_t} \,,$$

where

$$M_t \equiv \frac{\left(\frac{1}{\sigma}\right)(\hat{Y}_{t\wedge T_0})}{\left(\frac{1}{\sigma}\right)(a)}$$

is a martingale (with respect to $\hat{P}_a$). Together with (5.75) we know, M_t is a martingale under $\hat{P}_a$ if and only if $M_t D_t$ is a martingale under ${}^{-\lambda}P_a^{(4-\delta)}$, that is $((\frac{1}{\sigma})g)(Y_t)e^{-2\lambda t}$ is a martingale under ${}^{-\lambda}P_a^{(4-\delta)}$. □

We deduce $A^*(g) = 2\lambda g$ and $A^*(\frac{g}{\sigma}) = 2\lambda(\frac{g}{\sigma})$. From

$$A^*\left(\frac{g}{\sigma}\right) = gA^*\left(\frac{1}{\sigma}\right) + \frac{1}{\sigma}A^*(g) - \frac{\sigma'}{\sigma^2}g'$$

we obtain with $\frac{1}{\sigma}A^*(g) = 2\lambda\frac{g}{\sigma} = A^*(\frac{g}{\sigma})$

$$gA^*\left(\frac{1}{\sigma}\right) = \frac{\sigma'}{\sigma^2}g' \,. \tag{5.77}$$

By Itô's lemma we have

$$A(f(u)) = \frac{1}{2}f''(u)(u')^2 + f'(u)A(u)$$

and hence

$$A^*\left(\frac{1}{\sigma}\right) = \frac{(\sigma')^2}{\sigma^3} - \frac{1}{\sigma^2}A^*(\sigma)\,.$$

Together with (5.77) we have

$$\frac{\sigma'}{\sigma} - \frac{1}{\sigma'}A^*(\sigma) = \frac{g'}{g}\,,$$

from which we obtain

$$g(r) = c\,\frac{\sigma(r)}{\sqrt{\sigma'(r)}}\,r^{-\frac{3-\delta}{2}}\,e^{\frac{\lambda}{2}r^2}\,,$$

with a constant $c > 0$. Our aim is to derive (5.76), that is,

$$\frac{A^*(\sigma)}{\sigma'} = \frac{3-\delta}{2r} - \frac{\delta-1}{2r} - 2\lambda r\,,$$

and since $A^*(\sigma) = (\frac{3-\delta}{2r} - \lambda r)\sigma' + \frac{1}{2}\sigma''$, it remains to show

$$\frac{1}{2}\frac{\sigma''}{\sigma'} = -\frac{\delta-1}{2r} - \lambda r\,. \tag{5.78}$$

Remember, σ is the scale function of (R_t). Let b^R denote the drift of (R_t)

$$b^R(r) = \frac{\delta-1}{2r} + \lambda r\,,$$

then $b^R\sigma' + \frac{1}{2}\sigma'' = 0$ and hence (5.78).

5.5.3 The Three-Parameters-Family of Processes with Law ${}^{\lambda}P^{\delta,\mu}$

Based on Theorem 7 (Elworthy–Li–Yor [26]) we introduce a three-parameters-family of processes. From (5.73) we deduce

$${}^{\lambda}E_0^{\delta}[F(Y_{L_a-t},\,t \le L_a)\,e^{-\frac{\mu^2}{2}L_a}]$$
$$= \frac{1}{g(a)}\,{}^{-\lambda}E_a^{(4-\delta)}[F(Y_t,\,t \le T_0)\,e^{-(2\lambda+\frac{\mu^2}{2})T_0}]\,,$$

where $\lambda \geq 0$, $\delta > 2$, $\mu \geq 0$ and $g(a) = {}^{-\lambda}E_a^{(4-\delta)}(e^{-2\lambda T_0})$, and with the notation $2\theta \equiv 2\lambda + \frac{\mu^2}{2}$ we have

$$\frac{1}{g_\theta(a)}\, {}^{-\lambda}E_a^{(4-\delta)}[F(Y_t,\, t \leq T_0)\, e^{-2\theta T_0}] \\ = \frac{g(a)}{g_\theta(a)}\, {}^{\lambda}E_0^{\delta}[F(Y_{L_a - t},\, t \leq L_a)\, e^{-\frac{\mu^2}{2} L_a}]\,,$$

where $g_\theta(a) = {}^{-\lambda}E_a^{(4-\delta)}(e^{-2\theta T_0})$. This motivates us to define a three-parameters-family of diffusion processes with laws ${}^{\lambda}P_0^{\delta,\mu}$ as

$${}^{\lambda}P_0^{\delta,\mu}\big|_{\mathcal{F}_t} = \frac{\psi(Y_t)}{\psi(0)}\, e^{-\frac{\mu^2}{2}t} \cdot {}^{\lambda}P_0^{\delta}\big|_{\mathcal{F}_t}\,,$$

or more generally for arbitrary $x \geq 0$

$${}^{\lambda}P_x^{\delta,\mu}\big|_{\mathcal{F}_t} = \frac{\psi(Y_t)}{\psi(x)}\, e^{-\frac{\mu^2}{2}t} \cdot {}^{\lambda}P_x^{\delta}\big|_{\mathcal{F}_t} \tag{5.79}$$

for some increasing function ψ and $A\psi = \frac{\mu^2}{2}\psi$, $\psi \in C^2$, where A is the infinitesimal generator of ${}^{\lambda}P_x^{\delta}$. Indeed, with Itô's formula we see that $\frac{\psi(Y_t)}{\psi(x)}\, e^{-\frac{\mu^2}{2}t}$ is a martingale with respect to ${}^{\lambda}P_x^{\delta}$ and $(\mathcal{F}_t)$. Note that only with $f(t) \equiv e^{-ct}$ is $\psi(Y_t)f(t)$ a ${}^{\lambda}P_x^{\delta}$-martingale, since $\psi(Y_t)f(t)$ is a ${}^{\lambda}P_x^{\delta}$-martingale if and only if

$$L\psi(x)f(t) + \psi(x)f'(t) = 0,$$

where L denotes the infinitesimal generator of the process under ${}^{\lambda}P_x^{\delta}$, and hence $f'(t) = cf(t)$ for some constant c. Therefore, introducing an additional parameter μ is only reasonable in the above way (5.79).

First, let us consider the case $\lambda = 0$ where we simply write $P_x^{\delta,\mu}$ and P_x^{δ} instead of ${}^{0}P_x^{\delta,\mu}$ and ${}^{0}P_x^{\delta}$,

$$P_x^{\delta,\mu}\big|_{\mathcal{F}_t} = \frac{\psi(Y_t)}{\psi(x)}\, e^{-\frac{\mu^2}{2}t} \cdot P_x^{\delta}\big|_{\mathcal{F}_t}\,. \tag{5.80}$$

By means of the optional stopping theorem (see Revuz–Yor [89] § 2.3, p. 65) we obtain

$$\frac{\psi(a)}{\psi(x)} E_x^{\delta}[e^{-\frac{\mu^2}{2}T_a}] = 1\,. \tag{5.81}$$

For $\nu = \frac{\delta}{2} - 1 > -1$, $\infty > a > x > 0$ and $\mu > 0$ we know (see e.g. Kent [54] Theorem 3.1, Pitman–Yor [85] Proposition 2.3)

$$E_x^\delta[e^{-\frac{\mu^2}{2}T_a}] = \left(\frac{a}{x}\right)^\nu \frac{I_\nu(\mu x)}{I_\nu(\mu a)}\,,$$

thus we have

$$\psi(a) = \frac{I_\nu(\mu a)}{a^\nu}\,, \tag{5.82}$$

and

$$\psi(0) = \lim_{\varepsilon \downarrow 0} \frac{I_\nu(\mu\varepsilon)}{\varepsilon^\nu} = \left(\frac{\mu}{2}\right)^\nu \Gamma^{-1}(\nu+1)\,. \tag{5.83}$$

$(P_x^{\delta,\mu})$ is the family of laws of a diffusion with infinitesimal generator

$$\frac{1}{2}D^2 + \left(\frac{\psi'}{\psi} + \frac{\delta-1}{2y}\right) D$$

Watanabe [97] calls a diffusion process determined by such an infinitesimal generator a *Bessel diffusion process in the wide sense with index* $(\delta, \frac{\mu^2}{2})$. These diffusions arise as the Euclidean norm of a δ-dimensional Brownian motion with vector drift starting in zero. Denoting the norm of δ-dimensional Brownian motion with drift starting *at a nonzero vector* by R_t, we remark that R_t is not a Markov process, whereas $(R_t, \int_0^t \frac{ds}{R_s^2})$ has the Markov property (see Pitman–Yor [85, 86] or Rogers–Pitman [91]). From (5.82) we have

$$\frac{\psi'}{\psi}(y) = \mu \frac{I_{\nu+1}(\mu y)}{I_\nu(\mu y)}\,. \tag{5.84}$$

For example, in the case $\delta = 3$ we obtain by (5.82)

$$\frac{\psi(a)}{\psi(x)} = \frac{\sinh(\mu a)\, x}{a\, \sinh(\mu x)} \quad \text{and} \quad \frac{\psi(a)}{\psi(0)} = \frac{\sinh(\mu a)}{\mu a}\,. \tag{5.85}$$

Hence, $(P^{3,\mu})$ is the family of laws of a diffusion with infinitesimal generator

$$\frac{1}{2}D^2 + \frac{1}{y}D + (\mu \coth(\mu y) - \frac{1}{y})D = \frac{1}{2}D^2 + \mu \coth(\mu y) D\,.$$

This diffusion is the radial part of hyperbolic Brownian motion in dimension 3 (see Rogers–Pitman [91], Gruet [41]). As will be discussed

below, the reversed process at L_a is Brownian motion with drift $(-\mu)$ killed the first time it reaches zero.

Furthermore, we know, see Yor [102] p. 55,

Theorem 8 *Consider two transient diffusions with laws P_x and Q_x such that*

$$Q_x|_{\mathcal{F}_t} = D_t \cdot P_x|_{\mathcal{F}_t} \ .$$

Then

$$Q_x|_{\mathcal{F}_{L_a}} = (h(a)D_{L_a}) \cdot P_x|_{\mathcal{F}_{L_a}} \ ,$$

where $h(a) = \frac{\sigma'(a)}{\sigma(a)} \frac{s(a)}{s'(a)} \frac{\alpha(a)}{\beta(a)}$; σ and s denote the respective scale functions for Q_x and P_x and α and β are the respective diffusion coefficients for Q_x and P_x.

Continuing our discussion we apply Theorem 8 to (5.80). With

$$D_t \equiv \frac{\psi(Y_t)}{\psi(x)} \, e^{-\frac{\mu^2}{2}t}$$

we deduce from

$$h(a)\frac{\psi(a)}{\psi(x)} E_x^\delta(e^{-\frac{\mu^2}{2}L_a}) = 1 \tag{5.86}$$

the equality

$$E_x^{\delta,\mu}[F(Y_{L_a - t},\, t \le L_a)] \\ = (E_x^\delta(e^{-\frac{\mu^2}{2}L_a}))^{-1} \, E_x^\delta[F(Y_{L_a-t},\, t \le L_a)\, e^{-\frac{\mu^2}{2}L_a}] \,,$$

where for $\delta > 2$ arbitrary and for $x > 0$

$$E_x^\delta(e^{-\frac{\mu^2}{2}L_a}) = 2\nu \left(\frac{a}{x}\right)^\nu I_\nu(\mu(x \wedge a))K_\nu(\mu(x \vee a)) \,,$$

see e.g. Pitman–Yor [85], p. 330, (7.ẽ), and where (see (5.17))

$$E_0^\delta(e^{-\frac{\mu^2}{2}L_a}) = \frac{(\mu a)^\nu K_\nu(\mu\, a)}{2^{\nu-1}\Gamma(\nu)} \,.$$

Note that by means of (5.86) we obtain for $x = a$

$$E_a^\delta(e^{-\frac{\mu^2}{2}L_a}) = (h(a))^{-1} \,.$$

For instance, for the case $\delta = 3$ and $x = 0$ we have

$$D_t \equiv \left(\frac{\sinh(\mu Y_t)}{\mu Y_t} \right) e^{-\frac{\mu^2}{2} t} ,$$

and obtain

$$E_0^{3,\mu}[F(Y_{L_a - t},\, t \le L_a)] = e^{\mu a}\, E_0^3[F(Y_{L_a - t},\, t \le L_a)\, e^{-\frac{\mu^2}{2} L_a}] ,$$

since, see also (5.20),

$$E_0^3(e^{-\frac{\mu^2}{2} L_a}) = e^{-\mu a} .$$

Now let us look at the diffusion process with law $P_0^{3,\mu}$ time-reversed at L_a. As we remarked above it will turn out that this is Brownian motion with drift. Let P_x^0 denote the law of Brownian motion and P_x^3 the law of a 3-dimensional Bessel process starting in $x > 0$. From Doob's h-transform (5.71) we know

$$P_x^3\big|_{\mathcal{F}_t} = \frac{X_{t \wedge T_0}}{x} \cdot P_x^0\big|_{\mathcal{F}_t} ,$$

and together with (5.80) and (5.85) we have

$$P_x^{3,\mu}\big|_{\mathcal{F}_t} = \frac{\sinh(\mu X_{t \wedge T_0})}{\sinh(\mu x)}\, e^{-\frac{\mu^2}{2}(t \wedge T_0)} \cdot P_x^0\big|_{\mathcal{F}_t} . \tag{5.87}$$

Let $P_x^{(0)(\nu)}$ denote the law of Brownian motion with drift $\nu < 0$. Its scale function s is

$$s(x) = \frac{1}{2\nu}(1 - e^{-2\nu x}) .$$

Via Doob's h-transform (5.71) we have

$$P_x^{(\uparrow)(\nu)}\Big|_{\mathcal{F}_t} = \frac{e^{-2\nu X_{t \wedge T_0}} - 1}{e^{-2\nu x} - 1} \cdot P_x^{(0)(\nu)}\Big|_{\mathcal{F}_t} .$$

The Doob's h-transform (5.70) of Brownian motion with $h(x, t; \nu) \equiv e^{\nu x - \frac{\nu^2}{2} t}$ is identical in law to Brownian motion with drift ν, see e.g. Revuz–Yor [89], VIII.3, p. 327, or Borodin–Salminen [10], IV.28, that is

$$P_x^{(0)(\nu)}\Big|_{\mathcal{F}_t} = \frac{e^{\nu X_{t \wedge T_0} - \frac{\nu^2}{2}(t \wedge T_0)}}{e^{\nu x}} \cdot P_x^{(0)}\Big|_{\mathcal{F}_t} . \tag{5.88}$$

And finally, since

$$\frac{\sinh(\mu X_{t\wedge T_0})}{\sinh(\mu x)}\, e^{-\frac{\mu^2}{2}(t\wedge T_0)} = \frac{e^{2\mu X_{t\wedge T_0}}-1}{e^{2\mu x}-1}\, \frac{e^{-\mu X_{t\wedge T_0}-\frac{\mu^2}{2}(t\wedge T_0)}}{e^{-\mu x}}\,,$$

we obtain from (5.87) and from (5.88) with $\nu = -\mu$

$$P_x^{3,\mu}\big|_{\mathcal{F}_t} = \frac{e^{2\mu X_{t\wedge T_0}}-1}{e^{2\mu x}-1}\cdot P_x^{(0)(-\mu)}\Big|_{\mathcal{F}_t}\,,$$

this means, a diffusion process with law $P^{3,\mu}$ time-reversed is a Brownian motion with drift $(-\mu)$ killed the first time it reaches zero.

So far, we treated the case $\lambda = 0$ in full generality and as an application looked at the case $\delta = 3$. Now let us consider the general case $\lambda \geq 0$, i.e. we want to determine a ${}^{\lambda}P_x^{\delta}$-martingale D_t with

$${}^{\lambda}P_x^{\delta,\mu}\big|_{\mathcal{F}_t} = D_t \cdot {}^{\lambda}P_x^{\delta}\big|_{\mathcal{F}_t}\,.$$

We know for $\lambda = 0$ (see (5.80))

$$P_x^{\delta,\mu}\big|_{\mathcal{F}_t} = \frac{\psi(Y_t)}{\psi(x)}\, e^{-\frac{\mu^2}{2}t}\cdot P_x^{\delta}\big|_{\mathcal{F}_t}\,, \tag{5.89}$$

with ψ in (5.82) and via Girsanov's theorem we know

$${}^{\lambda}P_x^{\delta}\big|_{\mathcal{F}_t} = \exp\left\{\int_0^t \lambda Y_s dW_s - \frac{\lambda^2}{2}\int_0^t Y_s^2 ds\right\}\cdot P_x^{\delta}\big|_{\mathcal{F}_t}\,. \tag{5.90}$$

In order to obtain D_t, first we write the P_x^{δ}-martingale $\frac{\psi(Y_t)}{\psi(x)}\, e^{-\frac{\mu^2}{2}t}$ in (5.89) in a different form. We will tackle this problem more generally.

Let us assume, that H_t is a continuous, strictly positive semimartingale, and A_t is a continuous process with bounded variation such that $H_t e^{-A_t}$ is a local martingale. We will write $H_t e^{-A_t}$ in a different way. The continuous semimartingale H_t can be decomposed uniquely

$$H_t = m_t + a_t\,, \tag{5.91}$$

where m is a continuous local martingale and a a continuous adapted process of finite variation. With Itô's lemma we obtain

$$\begin{aligned} H_t e^{-A_t} &= \exp\{\ln H_t - A_t\} \\ &= \exp\left\{\int_0^t \frac{dH_s}{H_s} - \frac{1}{2}\int_0^t \frac{d\langle H,H\rangle_s}{H_s^2} - A_t\right\}\,, \end{aligned}$$

where the process $\langle H, H\rangle$ is the quadratic variation of H. With (5.91) and since $\langle H, H\rangle = \langle m, m\rangle$ we obtain

$$H_t e^{-A_t} = \exp\left\{\int_0^t \frac{dm_s}{H_s} - \frac{1}{2}\int_0^t \frac{d\langle m,m\rangle_s}{H_s^2} + \int_0^t \frac{da_s}{H_s} - A_t\right\}.$$

Since $H_t e^{-A_t}$ is a martingale, or equivalently

$$\int_0^t e^{-A_s} da_s - \int_0^t \frac{dA_s}{ds} e^{-A_s} H_s ds = 0$$

(see also the proof of Lemma 6), finally we have

$$H_t e^{-A_t} = \exp\left\{\int_0^t \frac{dm_s}{H_s} - \frac{1}{2}\int_0^t \frac{d\langle m,m\rangle_s}{H_s^2}\right\}. \tag{5.92}$$

Denoting $M_t \equiv \int_0^t \frac{dm_s}{H_s}$, equation (5.92) has the simple form

$$H_t e^{-A_t} = \exp\left\{M_t - \frac{1}{2}\langle M, M\rangle_t\right\}. \tag{5.93}$$

For example in the case $\delta = 3$, $x = 0$, we obtain

$$\frac{\sinh(\mu Y_t)}{\mu Y_t}\, e^{-\frac{\mu^2}{2}t} = \exp\left\{\int_0^t (\mu\coth(\mu Y_s) - \frac{1}{Y_s})dW_s \right.$$
$$\left. -\frac{1}{2}\int_0^t (\mu\coth(\mu Y_s) - \frac{1}{Y_s})^2 ds\right\}.$$

With the notation $N_t \equiv \int_0^t \lambda Y_s dW_s$, equation (5.90) can be written as

$${}^{\lambda}P_x^{\delta}\big|_{\mathcal{F}_t} = \exp\left\{N_t - \frac{1}{2}\langle N, N\rangle_t\right\} \cdot P_x^{\delta}\big|_{\mathcal{F}_t},$$

and hence, together with (5.93) we have

$$D_t = \exp\left\{M_t + N_t - \frac{1}{2}\langle M+N, M+N\rangle_t\right\}.$$

In the case $\delta = 3$, $x = 0$, we obtain

$$D_t = \exp\left\{\int_0^t (\mu\coth(\mu Y_s) - \frac{1}{Y_s} + \lambda Y_s)dW_s \right.$$
$$\left. -\frac{1}{2}\int_0^t (\mu\coth(\mu Y_s) - \frac{1}{Y_s} + \lambda Y_s)^2 ds\right\},$$

and the infinitesimal generator of the diffusion process under ${}^{\lambda}P_x^{3,\mu}$ is

$$\frac{1}{2}D^2 + (\mu \coth(\mu y) + \lambda y)\, D\,.$$

For arbitrary $\delta > 0$ analogous calculations together with (5.84) lead to

$$D_t = \exp\left\{ \int_0^t (\mu \frac{I_{\nu+1}(\mu Y_s)}{I_\nu(\mu Y_s)} + \lambda Y_s)\, dW_s \right. \\ \left. -\frac{1}{2}\int_0^t (\mu \frac{I_{\nu+1}(\mu Y_s)}{I_\nu(\mu Y_s)} + \lambda Y_s)^2\, ds \right\}\,,$$

and the infinitesimal generator of the diffusion process under ${}^{\lambda}P_x^{\delta,\mu}$ is

$$\frac{1}{2}D^2 + \left(\mu \frac{I_{\nu+1}(\mu y)}{I_\nu(\mu y)} + \frac{\delta - 1}{2y} + \lambda y \right) D\,,$$

and hence, we described the three-parameters-family of diffusion processes with law ${}^{\lambda}P^{\delta,\mu}$ in full generality. We remark that this family of diffusion processes is closely related with Ornstein–Uhlenbeck processes built from Brownian motion with drift.

Bibliography

[1] Ball, C.A. A review of stochastic volatility models with application to option pricing. *Financial Markets, Institutions and Instruments*, 2(5):55–71, 1993.

[2] Basawa, I.V. and Prakasa Rao, B.L.S. *Statistical Inference for Stochastic Processes*. Academic Press Inc., 1980.

[3] Bibby, B.M. Optimal combination of martingale estimating functions for discretely observed diffusion processes. Research report No.298, Dept. Theor. Statist., University of Aarhus, 1994.

[4] Bibby, B.M. A two-compartment model with additive white noise. Research report No.290, Dept. Theor. Statist., University of Aarhus, 1994.

[5] Bibby, B.M. Analysis of a tracer experiment based on compartmental diffusion models. Research report No.303, Dept. Theor. Statist., University of Aarhus, 1995.

[6] Bibby, B.M. On estimation in compartmental diffusion models. Research report No.305, Dept. Theor. Statist., University of Aarhus, 1995.

[7] Bibby, B.M and Sørensen, M. Martingale estimation functions for discretely observed diffusion processes. *Bernoulli*, 1(1/2):17–39, 1995.

[8] Bibby, B.M and Sørensen, M. On estimation for discretely observed diffusions: A review. Research report No.334, Dept. Theor. Statist., University of Aarhus, 1995.

[9] Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

[10] Borodin, A.N. and Salminen, P. *Handbook of Brownian Motion – Facts and Formulae.* Probability and Its Applications. Birkhäuser Verlag Basel Boston Berlin, 1996.

[11] Breiman, L. First exit times from a square root boundary. *Fifth Berkeley Symposium*, 2(2):9–16, 1967.

[12] Breiman, L. *Probability.* Addison-Wesley Publishing Company, 1968.

[13] Brockwell, P.J. and Davis, R.A. *Time Series: Theory and Methods.* Springer-Verlag, New York, 1987.

[14] Chen, R. and Scott, L. Pricing interest rate options in a two-factor Cox–Ingersoll–Ross model of the term structure. *The review of financial studies*, 5(4):613–636, 1992.

[15] Chesney, M., Geman, H., Jeanblanc-Picqué, M., and Yor, M. Some combinations of Asian, Parisian and Barrier options. In Dempster, M.A.H and Pliska, S.R., editors, *Mathematics of Derivative Securities*, Publications of the Newton Institute. Cambridge University Press, 1997.

[16] Cox, J.C., Ingersoll, J.E. Jr., and Ross, S.A. A theory of the term structure of interest rates. *Econometrica*, 53(2), 1985.

[17] Cramér, H. *Mathematical Methods of Statistics.* Princeton University Press, 1946.

[18] Dacunha-Castelle, D. and Florens-Zmirou, D. Estimation of the coefficients of a diffusion from discrete observations. *Stochastics*, 19:263–284, 1986.

[19] Deelstra, G. and Parker, G. A covariance equivalent discretisation of the CIR model. *Proceedings of the 5th AFIR International Colloquium*, 2:731–747, 1995.

[20] Delbaen, F. Consols in the CIR model. *Mathematical Finance*, 3(2):125–134, 1993.

[21] Delbaen, F. and Shirakawa, H. Option pricing for constant elasticity of variance model. Preprint, 1996.

[22] Delbaen, F. and Shirakawa, H. Squared bessel processes and their applications for the square root interest rate model. Preprint, 1996.

[23] Dellacherie, C., Maisonneuve, B., and Meyer, P.-A. *Processus de Markov (fin); Complements de calcul stochastique*, volume 5 of *Probabilites et potentiel*. Publications de l'Institut de mathématique de l'Université de Strasbourg, 1992.

[24] Doob, J.L. *Classical Potential Theory and Its Probabilistic Counterpart*. Springer-Verlag New York, 1984.

[25] Duffie, D. and Singleton, K.J. An econometric model of the term structure of interest rate swap yields. Preprint, Graduate School of Business, Stanford University, 1995.

[26] Elworthy, K.D., Li, X.M., and Yor, M. Two papers on: Strictly local martingales. 2.: The importance of strictly local martingales: Applications to radial Ornstein-Uhlenbeck processes. Prépublication 346, Laboratoire de Probabilités de l'Université Paris VI, 1996.

[27] Embrechts, P., Klüppelberg, C., and Mikosch, T. *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag Berlin, 1997.

[28] Embrechts, P., Samorodnitsky, G., Dacorogna, M.M., and Müller, U.A. How heavy are the tails of a stationary HARCH(k) process? A study of moments. Preprint, Olsen & Associates, Zürich, 1996.

[29] Engle, R.F. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica*, 50:987–1008, 1982.

[30] Feigin, P.D. Maximum likelihood estimation for continuous-time stochastic processes. *Adv. Appl. Probab.*, 8:712–736, 1976.

[31] Florens-Zmirou, D. Approximate discrete-time schemes for statistics of diffusion processes. *Statistics*, 20:547–557, 1989.

[32] Fournie, E. and Talay, D. Application de la statistique des diffusions a un modèle de taux d'intérêt. Technical report, INRIA, 1996. to appear in Finance.

[33] Frydman, H. Asymptotic inference for the parameters of a discrete-time square-root process. *Mathematical Finance*, 4(2): 169–181, 1994.

[34] Geman, H. and Yor, M. Bessel processes, Asian options, and perpetuities. *Mathematical Finance*, 3(4):349–375, 1993.

[35] Geman, H. and Yor, M. Pricing and hedging double-barrier options: A probabilistic approach. *Mathematical Finance*, 6(4):365–378, 1996.

[36] Getoor, R.K. The Brownian escape process. *Ann. Prob.*, 7(5):864–867, 1979.

[37] Getoor, R.K. and Sharpe, M.J. Excursions of Brownian motion and Bessel processes. *Z. Wahrsch. Verw. Gebiete*, 47:83–106, 1979.

[38] Gihman, I.I. and Skorokhod, A.V. *Stochastic Differential Equations.* Springer-Verlag, Berlin, 1972.

[39] Godambe, V.P. and Heyde, C.C. Quasi-likelihood and optimal estimation. *Int. Statist. Rev.*, 55:231–244, 1987.

[40] Göing, A. Estimation in financial models. Risklab report Dec. 1995, ETH Zürich, 1995.

[41] Gruet, J.-C. Semi-groupe du mouvement brownien hyperbolique. *Stochastics and Stochastics Reports*, 56:53–61, 1996.

[42] Heyde, C.C. On combining quasi-likelihood estimating functions. *Stochastic Processes and their Applications*, 25:281–287, 1987.

[43] Heyde, C.C. *Quasi-Likelihood and its Application.* Springer Series in Statistics. Springer-Verlag New York Berlin Heidelberg, 1997.

[44] Honoré, P. Estimation of a dynamic one factor continuous-time term-structure model. Department of Finance, The Aarhus School of Business, 1996.

[45] Hutton, J.E. and Nelson, P.I. Quasi-likelihood estimation for semimartingales. *Stochastic Processes and their Applications*, 22:245–257, 1986.

[46] Ibragimov, I.A. and Has'minskii, R.Z. *Statistical Estimation.* Springer-Verlag, New York, 1981.

[47] Ikeda, N. and Watanabe, S. *Stochastic Differential Equations and Diffusion Processes.* North-Holland Publishing Company, 1981.

[48] Ismail, M.E.H. and Kelker, D.H. The Bessel polynomials and the Student t distribution. *SIAM J. on Math. Anal.*, 7(1):82–91, 1976.

[49] Itô, K. and McKean Jr., H.P. *Diffusion Processes and their Sample Paths.* Springer-Verlag Berlin Heidelberg, reprint of the 1974 edition, 1996.

[50] Jacod, J. Statistics of diffusion processes: Some elements. Technical Report 94.17, Istituto per le Applicazioni della Matematica e dell'Informatica, 1994.

[51] Jeulin, T. *Semi-Martingales et Grossissement d'une Filtration.* LNM 833. Springer-Verlag Berlin Heidelberg, 1980.

[52] Kallianpur, G. *Stochastic Filtering Theory.* Springer-Verlag, New York, 1980.

[53] Karlin, S. and Taylor, H.M. *A Second Course in Stochastic Processes.* Academic Press, 1981.

[54] Kent, J. Some probabilistic properties of Bessel functions. *Ann. Prob.*, 6(5):760–770, 1978.

[55] Kessler, M. Simple and explicit estimating functions for a discretely observed diffusion process. Research report No.336, Dept. Theor. Statist., University of Aarhus, 1996.

[56] Kessler, M. Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.*, 24(2):211–229, 1997.

[57] Kessler, M. and Sørensen, M. Estimating equations based on eigenfunctions for a discretely observed diffusion process. Research report No.332, Dept. Theor. Statist., University of Aarhus, 1995.

[58] Kloeden, P.E. and Platen, E. *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag, New York, 1992.

[59] Kloeden, P.E., Schurz, H., Platen, E., and Sørensen, M. On effects of discretization on estimators of drift parameters for diffusion processes. *J. Appl. Probab.*, 33(4):1061–1076, 1996.

[60] Kunita, H. Absolute continuity of Markov processes and generators. *Nagoya Math. J.*, 36:1–26, 1969.

[61] Kunita, H. *Stochastic Flows and Stochastic Differential Equations*, volume 24 of *Cambridge studies in advanced mathematics*. Cambridge University Press, 1990.

[62] Kutoyants, Yu.A. *Parameter Estimation for Stochastic Processes*. Heldermann-Verlag, Berlin, 1984.

[63] Lamberton, D. and Lapeyre, B. *Introduction to Stochastic Calculus Applied to Finance*. Chapman & Hall London, 1996.

[64] Lamperti, J. Semi-stable Markov processes I. *Z. Wahrsch. Verw. Gebiete*, 22:205–225, 1972.

[65] Lebedev, N.N. *Special Functions and their Applications*. Dover Publications, New York, 1972. Translated and edited by Richard A. Silverman.

[66] Leblanc, B. *Modélisations de la Volatilité d'un Actif Financier et Applications*. PhD thesis, Université de Paris VII, 1997.

[67] Lindgren, B.W. *Statistical Theory*. Macmillan Publishing Co., 3rd edition, 1976.

[68] Liptser, R.S. and Shiryaev, A.N. *Statistics of Random Processes I, II*. Springer-Verlag, New York, 1977.

[69] McLeish, D.L. and Kolkiewicz, A.W. Fitting diffusion models in finance. Preprint.

[70] Millar, P.W. Random times and decomposition theorems. *Proc. Symp. Pure Mathematics*, 31:91–103, 1977.

[71] Musiela, M. and Rutkowski, M. *Martingale Methods in Financial Modelling*. Springer-Verlag Berlin, 1997.

[72] Nagasawa, M. Time reversion of Markov processes. *Nagoya Mathematical Journal*, 24:177–204, 1964.

[73] Nagasawa, M. *Schrödinger Equations and Diffusion Theory*, volume 86 of *Monographs in Mathematics*. Birkhäuser Verlag Basel Boston Berlin, 1993.

[74] Novikov, A.A. A martingale approach to first passage problems and a new condition for Wald's identity. In *Lecture Notes in Control and Information Sci.*, Volume 36 of *Stochastic differential systems (Visegrád, 1980),* 146–156. Springer-Verlag Berlin New York, 1981.

[75] Novikov, A.A. A martingale approach in problems on first crossing time of nonlinear boundaries. *Proc. of the Steklov Institute of Mathematics*, 4:141–163, 1983.

[76] Novikov, A.A. Limit theorems for the first passage time of autoregression process over a level. *Proc. of the Steklov Institute of Mathematics*, 4:169–186, 1994.

[77] Øksendal, B. *Stochastic Differential Equations*. Springer-Verlag, Berlin, 2nd edition, 1989.

[78] Overbeck, L. and Rydén, T. Estimation in the Cox–Ingersoll–Ross model. *Econometric Theory*, 13(3):430–461, 1997.

[79] Pedersen, A.R. Maximum likelihood estimation based on incomplete observations for a class of discrete time stochastic processes by means of the Kalman filter. Research Report No.272, Dept. Theor. Statist., University of Aarhus, 1993.

[80] Pedersen, A.R. Quasi-likelihood inference for discretely observed diffusion processes. Research Report No.295, Dept. Theor. Statist., University of Aarhus, 1994.

[81] Pedersen, A.R. Statistical analysis of Gaussian diffusion processes based on incomplete discrete observations. Research Report No.297, Dept. Theor. Statist., University of Aarhus, 1994.

[82] Pedersen, A.R. Uniform residuals for discretely observed diffusion processes. Research Report No.292, Dept. Theor. Statist., University of Aarhus, 1994.

[83] Pedersen, A.R. Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes. *Bernoulli*, 1(3):257–279, 1995.

[84] Pedersen, A.R. A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scand. J. Statist.*, 22(1):55–71, 1995.

[85] Pitman, J. and Yor, M. Bessel processes and infinitely divisible laws. In *Stochastic Integrals*, Volume 851 of *Lecture Notes in Mathematics*, 285–370. Springer-Verlag Berlin Heidelberg, 1980.

[86] Pitman, J. and Yor, M. Inversion du temps et processus de Bessel. Working paper, 1981.

[87] Pitman, J. and Yor, M. A decomposition of Bessel bridges. *Z. Wahrsch. Verw. Gebiete*, 59:425–457, 1982.

[88] Pitman, J. and Yor, M. Sur une décomposition des ponts de Bessel. In *Functional Analysis in Markov Processes*, Volume 923 of *Lecture Notes in Mathematics*, 276–285. Springer-Verlag Berlin Heidelberg, 1982.

[89] Revuz. D. and Yor, M. *Continuous Martingales and Brownian Motion.* Springer-Verlag Berlin Heidelberg, 1991.

[90] Rogers, L.C.G. Which model for term-structure of interest rates should one use? *The IMA Volumes in Mathematics and its Applications, Springer-Verlag New York*, 65:93–115, 1995.

[91] Rogers, L.C.G. and Pitman, J.W. Markov functions. *Ann. Prob.*, 9(4):573–582, 1981.

[92] Rogers, L.C.G. and Williams, D. *Diffusions, Markov Processes and Martingales,* Volume 2: Itô Calculus. Wiley, Chichester, 1987.

[93] Sharpe, M.J. Some transformations of diffusions by time reversal. *Ann. Prob.*, 8(6):1157–1162, 1980.

[94] Shepp, L. First passage problem for the Wiener process. *Ann. Math. Stat.*, 38:1912–1914, 1967.

[95] Shiga, T. and Watanabe, S. Bessel diffusions as a one-parameter family of diffusion processes. *Z. Wahrsch. Verw. Gebiete*, 27:37–46, 1973.

[96] Szatzschneider, W. Bessel processes and pricing interest rate derivative securities in generalized CIR model. Universidad Anaáhuac, Mexico City.

[97] Watanabe, S. On time inversion of one-dimensional diffusion processes. *Z. Wahrsch. Verw. Gebiete*, 31:115–124, 1975.

[98] Widder, D. V. *The heat equation.* Academic Press, Inc., 1975.

[99] Williams, D. Path decomposition and continuity of local time for one-dimesional diffusions, I. *Proc. London Math. Soc. (3)*, 28:738–768, 1973.

[100] Wong, E. and Hajek, B. *Stochastic Processes in Engineering Systems.* Springer-Verlag, New York, 1985.

[101] Yor, M. On square-root boundaries for Bessel processes and pole-seeking Brownian motion. In *Stochastic Analysis and Applications*, Volume 1095 of *Lecture Notes in Mathematics*, 100–107. Springer-Verlag Berlin Heidelberg, 1984.

[102] Yor, M. *Local Times and Excursions for Brownian Motion.* Volume 1. 1995 of *Lecciones en Matemáticas.* Universidad Central de Venezuela, 1995.

[103] Yor, M. *Some Aspects of Brownian Motion, Part II: Some recent Martingale Problems.* Lectures in Mathematics, ETH Zürich. Birkhäuser Verlag Basel, 1997.