Stationarity in Markovian marked point process and jump-diffusion models

by Jerôme Baltzersen

August 3, 2010

Thesis adviser: Martin Jacobsen

Thesis for the Master degree in Mathematics.
Department of Mathematical Sciences, University of Copenhagen.

This thesis is printed using Computer Modern 11pt.
Layout and typografy by the author using LaTeX

Contents

Abstract

Many phenomena exhibit a piecewise continuous non-deterministic behavior. Jump-diffusions with jumps from a renewal process are an obvious choice for modeling such behavior. The key question treated in this thesis is to investigate when a stationary distribution exists and, if possible, to find it.

Two different models are considered. The first specifies the target just after each jump, while the second specifies the change relative to the position of the process just before the jump. Applications of both models are briefly touched upon.

The existence and uniqueness of a stationary distribution is shown using the key renewal theorem for a very broad class of models. In case of the jump target model, renewal theory offers a way to find the stationary distribution while a integro-differential equation, sometimes solvable, for the density is also presented.

The jump size model is harder to deal with. In certain cases it is possible to reduce the analysis to a jump target model. Nevertheless, explicitness is much harder to attain.

For both models, results are more tangible when the jumps stem from a Poisson process. For such models, in the jump target case we are able to find the stationary density for many popular diffusion choices.

Resume

Mange fænomener udviser en stykkevis ikke-deterministisk kontinuert opførsel. Et oplagt valg til at modellere disse er spring-diffusioner med spring fra en underliggende fornyelsesproces. Det centrale tema i specialet er at undersøge, hvornår en stationær fordeling eksisterer, og hvis muligt at finde den.

To forskellige modeller betragtes. Den første angiver den position, der springes til, mens den anden angiver springstørrelsen relativt til processens position umiddelbart inden springet. Anvendelser af begge modeller berøres kort.

Eksistensen of entydigheden af en stationær fordeling vises for en bred klasse af modeller via fornyelsessætningen. I modellen der angiver springpositionen giver fornyelses teori mulighed for at finde den stationære fordeling. Derudover præsenteres en integro-differentialligning for tætheden, der i specialtilfælde kan løses.

Modellen, der angiver springstørrelse, er vanskeligere at arbejde med. I en række tilfælde kan analysen reduceres til en springpositionsmodel. Ikke desto mindre er det langt sværere at opnåeksplicitte resultater.

For begge modeller bliver resultaterne mere håndgribelige, når springene stammer fra en Poisson-proces. I sådanne modeller kan den stationære fordeling findes for en række populære diffusioner i springpositionstilfældet.

Preface

> *'Well!' thought Alice to herself, 'after such a fall as this, I shall think nothing of tumbling down stairs!'*
> Lewis Carroll, "Alice's Adventures in Wonderland," 1865.

This is a bittersweet feeling. Writing this preface represents the conclusion of two things: i) the thesis itself and ii) my time as a student. Both have with their lows and highs been dear projects into which I have poured much energy and passion.

The sweetness: No undertaking can be truly evaluated while in the midst of it; how could you know if it's up for a happy ending? Wrapping something up means completing what one set out to do.

The bitterness: When having enjoyed the daily labor throughout, when realizing that the process too, not only the result, matters; when having met great people; when having seen beauty at a different scale; why would it have to stop?

Ahead lies uncertainty; opportunity!

Jerôme Baltzersen
Copenhagen, August 2010
mail@baltzersen.info

Acknowledgments

No written piece of a certain length sees the light of day without the involvement of others. Family, friends, and colleagues have all met me with great understanding for how supremely important it was for me to focus on writing the thesis. Their patience has allowed me to neglect daily life a bit while attending to the writing process.

An adviser is expected to play a fundamental role in the preparation of a student's thesis, but Martin Jacobsen has done much more. Our weekly meetings—even during his private vacation!—has kept my pace up. And, when close to despair, his soothing comments like "enlightenment comes in stages of confusion" has helped keep up the spirit as well.[1] Also, I've learned a bit 'bout tennis.

While the faculty staff has been supporting, willing to answer questions and discuss theories, Susanna Ditlevsen has involved herself with great enthusiasm. Her taking time from a busy workday and cheering passionately for even the smallest results and progress in analyzing her models has made it easy to stay motivated.

Morten Hornbech, a dear friend, has read the manuscript with utmost care and provided a myriad of most-welcome suggestions. His industrious work, I'm sure, makes the reading experience for all who follow more pleasant. In addition his help with the numerical analysis has been indispensable.

Finally, I would like to thank two persons playing a fundamental role throughout my student years, without whom studying would have been half as much fun. Anders Thorup who from day one took me under his wing, encouraged me, introduced me to new theories, and most importantly believed in whatever abilities I might have. Everyone needs somebody who believes in them. Ernst Hansen has patiently guided me through the not particularly linear path of my studies, and without him many of my ambitions and ideas would have been unmet. His excellent teaching has also been of great benefit to me; and the entire student body almost surely.

[1] My translation of "erkendelse kommer i stadier af forvirring."

Introduction

I have been very fortunate to be exposed to a vast amount of different mathematics, abstract as well as applied; technical as well as hands on. Most mathematicians know the feeling of immense beauty when an argument is of such clarity one could mistake it for being divine.[2] On the other hand building a mathematical model can be very intriguing solving a real problem and adding to our understanding of the world surrounding us. Describing a phenomenon mathematically such that its most outspoken characteristics are fulfilled allows for the powerful framework of mathematical analysis to investigate implications beyond the initial comprehension.

When passionate about both, how does one combine them? I think rigorous probabilistic modeling qualifies. The tools needed are sufficiently complex for a deep mathematical understanding of underlying concepts while at the same time, the very origin of probability theory dictates a high involvement with reality.

Therefore, in short, this thesis links my two main interests, while at the same time expanding my horizon of a mathematical understanding of the world.

How to read this thesis

This thesis contains three parts:

- discussion of preliminary theory,
- building and analysis of models, and
- applications of the models.

The vast majority of the theory needed for building the models was new to the author prior beginning the work on the thesis. Some of it is therefore developed, while some is quoted simply in order to get the notation and formulation of the results in a manner useful in this thesis.

The preliminary theory naturally splits into three sections: marked point processes, stochastic calculus, and renewal theory. The first topic was known to the author beforehand and presented briefly. In contrast, stochastic calculus as well as renewal theory were topics completely new to the author. Here the theory is developed, motivated, and exemplified to the extent allowed in the setting of a thesis.

The reader well-acquainted with all three parts may want simply to skim the introductory chapter to get a feel for the notation. The reader unfamiliar with marked point processes is most likely better off beginning with, or at least in parallel, reading a textbook such as [Ja06]. The chapters on stochastic calculus and renewal

[2] "You don't have to believe in God, but you should believe in The Book.", Paul Erdös

theory are, albeit superficial compared to a standard textbook treatment, more self-contained and should have somewhat of an isolated value.

We then proceed to build the actual models. Two similar models only differing in what happens at jump times are described. The jump target model, Chapter 2, specifies where the next process begins. The jump size model, Chapter 3, specifies what increment should be added to the pre-jump position of the process.

The reader is strongly encouraged to begin with the jump target model. Not only can more explicitness be obtained in this model, but it also serves as the foundation for the jump size model. That is, many of the results in the jump size model are derived by realizing that certain cases can be reduced to a jump target setting, which opens up for results already established.

In the jump target model quite a few nice results are established and the reader will find explicit examples of stationary densities and functional equations describing such. The jump target model is mature enough to be applied even though there are plenty of outstanding questions. Unfortunately this is nowhere near the case for the jump size model. It turns out to be dreadfully difficult to even establish the existence of the stationary distribution and in fact the author was not able to find a single case where this could be proved.

To ignite the creativity of the reader wishing to apply the models (probably primarily the jump target model) two such are presented in Chapter 4. The jump target model is applied to the potential of a neuron with some success and a strong indication of perspectives. A model of beta cells producing insulin are built using the jump size model, but with the model being so intrinsically inaccessible little ground is gained.

CHAPTER 1

Point processes, stochastic calculus, and renewal theory

The title reveals that this chapter contains the treatment of three theories each extensively covered by textbooks. It is a lot of ground to cover. However, knowing what the applications are in later chapters we may offer a concise picture of the involved theory often using a notation that highlights the important parts. Naturally, some generality is lost hereby for which the interested reader is encouraged to pick up the references.

While many books have served as inspiration some have been indispensable and reoccurring in the process of understanding the theory. The point process theory is based entirely on [Ja06]. The stochastic calculus needed in this thesis only uses the Brownian motion as integrators and as such [Ok00] is a fitting choice, while for a more detailed treatment of diffusions [Ja08], in particular chapters 9 and 11, is recommended. Renewal theory has many excellent expositions, but [KT75] has been the authors main source of inspiration. Some notation and results are found in Appendix A.

Classics such as [Fe66] and [KT75] are lucid standard references for probability theory and stochastic processes with good reason, while the author is trained using [Ha09] and [Ja03] providing a home court advantage.

The highlights of this chapter is the construction of piecewise deterministic Markov processes using marked point processes and the analysis of one-dimensional diffusions, particularly stationarity. Finally, the key renewal theorem and regenerative processes discussed are both essential for proving the existence of stationary distributions in the models later on.

The notation used throughout is rather standard and the well-acquainted reader should be able to go directly to the models of Chapters 2 and 3 using this chapter more as a reference.

1.1. General notation

By ":=" we mean "is defined to be equal to," $x \wedge y := \min\{x, y\}$, and $f(t) \sim g(t)$ is equivalent to $\lim_{t\to\infty} f(t)/g(t) = 1$. Certain function spaces are encountered throughout the thesis as well. $C^n(]a, b[)$ are the n times differentiable functions on the interval $]a, b[$, while $C_0^n(]a, b[)$ and $C_b^n(]a, b[)$ in addition requires compact

support and boundedness, respectively. Here $]a, b[$ can be substituted with any open subset of $\mathbb{R}$. Integration of f with respect to t is denoted $\int f\,dt$ or $\int dt f$ depending on what notation is more convenient in the given context.

Denote the abstract background space by $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{P}$ the probability measure. We are interested in real-valued stochastic variables mapping into $(\mathbb{R}, \mathcal{B})$, where $\mathcal{B}$ is the Borel-algebra on $\mathbb{R}$. We further write $\mathbb{R}_0 = [0, \infty[$ and $\overline{\mathbb{R}}_+ =]0, \infty]$ with corresponding σ-algebras $\mathcal{B}_0$ and $\overline{\mathcal{B}}_+$. The degenerate measure at $\{x\}$ is denoted by ε_x.

Let X be a real-valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. The transformed probability measure then becomes $X(\mathbb{P})$ and if X is $\mathbb{P}$-integrable

$$\mathrm{E}[X] := \int X d\mathbb{P}.$$

If X is strictly positive with distribution F it holds that

$$\mathrm{E}[X] = \int_0^\infty \int_0^x dt\; dF(x) = \int_0^\infty \int_t^\infty dF(x)\; dt = \int_0^\infty 1 - F(x)\; dx, \tag{1.1}$$

where the survivor function $\overline{\mathbb{P}}(x) := 1 - F(x)$ is often used. In the discrete case this becomes

$$\mathrm{E}[X] = \sum_{k=0}^{\infty} \mathbb{P}(X \geq k). \tag{1.2}$$

A stochastic process is a family of stochastic variables denoted by $\mathbf{X} = (X_t)_{t \in I}$, usually thinking of t belonging to the half line $I = [0, \infty[$, where each X_t maps into some general state space denoted by $(G, \mathcal{G})$. Often a probability field is equipped with an increasing sequence of sub-σ-algebras, that is $\mathcal{F}_s \subset \mathcal{F}_t$ whenever $s \leq t$, called a filtration and $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ is called a filtered probability space. A stochastic process X_t on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ is said to be $\mathcal{F}_t$-adapted—or simply $(X_t, \mathcal{F}_t)$ is adapted—if X_t is $\mathcal{F}_t$ measurable for all $t \geq 0$. The mapping from $(\mathbb{R}_0 \times \Omega, \mathcal{B}_0 \otimes \mathcal{F})$ into $(G, \mathcal{G})$ given by

$$(t, \omega) \mapsto X_t(\omega) \tag{1.3}$$

is considered. $\mathbf{X}$ is said to be measurable if this map is measurable in both coordinates. A stochastic process is predictable (or previsible) if, intuitively, its values are known just ahead of time. Formally, it is required that the map from (1.3) is measurable with respect to the sub-σ-algebra generated by

$$]s, \infty[\times F \qquad s \in \mathbb{R}_0, F \in \mathcal{F}_s.$$

The Brownian motion is denoted by $\mathbf{B} = (B_t)_{t \geq 0}$ and used throughout. It is defined in [Ja03] p. 132 and all of p. 131-158 is assumed well-known to the reader. A standard Brownian motion has drift 0, variance 1, and begins at 0.

Note, that ∂^n denotes derivatives in distributional sense (see Section A.1), while $\frac{d^n}{dx^n}$ is used for regular differentiation.

We also need a notion of convergence of a sequence of probability measures $(\nu_n)_1^\infty$. We write $\nu_n \xrightarrow{w} \nu_0$ and say that ν_n converges to ν_∞ if

$$\lim_{n \to \infty} \int f(y)\, d\nu_n(y) = \int f(y)\, d\nu_\infty$$

for all $f \in C_b(G)$. Continuity is only well-defined if G is equipped with a topology, which is always the case in this thesis as we restrict attention to the case $G = \mathbb{R}$.

1.2. Marked point processes

Intuitively, a marked point process (MPP) is a stochastic process jumping between a certain set of marks/labels called the mark space at discrete stochastic times. Albeit being simply to grasp, draw, and explain intuitively, the mathematical rigor required is rather extensive. A major part of this thesis is concerned with the connection between MPPs and piecewise deterministic Markov processes (PDMP), see Section 1.5.

Let $T_1, \dots$ denote the stochastic jump times and let $(E, \mathcal{E})$ be the mark space with the corresponding σ-algebra. Adjoin the irrelevant mark, ∇, used for an event never occurring, that is if $T_n = \infty$ for some n. Although this event never occurs it formally still needs a mark associated. Define $(\overline{E}, \overline{\mathcal{E}}) = (E \cup \{\nabla\}, \sigma(\mathcal{E}, \{\nabla\}))$.

Definition 1.1 (Marked point process)**.** A *marked point process* is the pair given by

$$((\mathcal{T}, \mathcal{Y}) = ((T_n)_{n\geq 1}, (Y_n)_{n\geq 1})$$

with $T_n \in \overline{\mathbb{R}}_+$ and $Y_n \in \overline{E}$ both random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. It must satisfy

$$\mathbb{P}(0 < T_1 \leq T_2 \leq \dots) = 1, \tag{1.4}$$
$$\mathbb{P}(T_n < T_{n+1}, T_n < \infty) = \mathbb{P}(T_n < \infty), \tag{1.5}$$
$$\mathbb{P}(\lim_{n\to\infty} T_n = \infty) = 1, \tag{1.6}$$
$$\mathbb{P}(Y_n \in E, T_n < \infty) = \mathbb{P}(T_n < \infty) \quad \text{for } n \geq 1, \text{ and} \tag{1.7}$$
$$\mathbb{P}(Y_n \in \{\nabla\}, T_n = \infty) = \mathbb{P}(T_n = \infty) \quad \text{for } n \geq 1. \tag{1.8}$$

Condition (1.4) through (1.6) only involves the jump times. When disregarding the marks a process satisfying these is called a Simple Point Process (SPP) and results concerning MPPs can be translated into results about SPPs. Condition (1.4) together with (1.5) ensures that as long as a jump occurs in finite time the jump times are strictly increasing with probability 1.

Condition (1.6) is of particular interest. It rules out so-called explosions, that is infinitely many jumps in finite time. The theory excluding explosions suffices for our applications albeit some results need the notion of MPPs with explosion; a class referred to by MPP_{ex} and used in Theorem 1.1.

The marks could be virtually anything: A real number, a tuple of parameters describing a function or even functions as long as the mark space is measurable. Nevertheless, this thesis restricts the mark spaces to Borel-spaces in order to assure that transition probabilities exist. Condition (1.8) formalizes the use of the irrelevant mark for the event never occurring. Note that the definition does not guarantee genuine jumps as it might be that $y_n = y_{n+1}$ for some n the only restriction being that both $y_n, y_{n+1} \in E$.

Rather than viewing, as Definition 1.1 suggests, an MPP as a tuple of $\overline{\mathbb{R}}_+$- and $\overline{E}$-valued random variable, it can be fruitful to impose additional structure. An MPP $(\mathcal{T}, \mathcal{Y})$ can be viewed as a random variable mapping into $(K_E, \mathcal{K}_E)$ where

$$\begin{aligned} K_E = \Big\{ ((t_n), (y_n)) \in \overline{\mathbb{R}}_+^{\mathbb{N}} \times \overline{E}^{\mathbb{N}} : \\ 0 < t_1 \leq \dots \leq t_n \leq \infty, t_k < t_{k+1} \text{ if } t_k < \infty; \\ y_n \in E \text{ if and only if } t_n < \infty \Big\} \end{aligned}$$

equipped with the σ-algebra $\mathcal{K}_E$ generated by the coordinate projections

$$T_n^\circ((t_k),(y_k)) = t_n \qquad \text{and} \qquad Y_n^\circ((t_k),(y_k)) = y_n.$$

The distribution of $(\mathcal{T},\mathcal{Y})$ on $(K_E,\mathcal{K}_E)$ is well-defined and given by $R = (\mathcal{T},\mathcal{Y})(\mathbb{P})$.

Consider for a brief moment a SPP; a point process where we only keep track of times. Sometimes it is more convenient to describe the SPP as a counting process N_t yielding the number of jumps before time t. This is a of course again a stochastic variable and it is not hard to imagine that using $T_n = \inf\{t \geq 0 : N_t = n\}$ the jump times can be recovered showing the description is bijective.

Analogously, it is sometimes advantageous to deal with an MPP on a similar space rather than $(K_E,\mathcal{K}_E)$. Adding the marks gives another dimension and therefore the appropriate question to pose is how many jumps occurred into a subset A of the marks. This is denoted by $N_t(A)$ and defined as

$$N_t(A) := \sum_{n=1}^{\infty} 1_{(T_n \leq t, Y_n \in A)}$$

with the auxiliary notation for all jumps during $]0,t]$

$$\overline{N}_t(A) = N_t(E).$$

The jump times are still easily recovered using

$$T_n = \inf\{t \geq 0 : \overline{N}_t = n\},$$

whereas it is less obvious how to, in a measurable fashion, recover the Y_n's. However, it holds for any $A \in \mathcal{E}$ that $\mathbb{P}$-a.s.

$$(Y_n \in A) = \bigcup_{K'=1}^{\infty} \bigcap_{K=K'}^{\infty} \bigcup_{k=1}^{\infty} \left(\overline{N}_{(k-1)/2^K} = n-1, N_{k/2^K}(A) - N_{(k-1)/2^K}(A) = 1\right)$$

as argued in [Ja06] p. 14-15. This is adequate for our needs, but it should be noted that a *random counting measure* ϱ is cleverly introduced such that $\varrho(]0,t] \times A) = N_t(A)$ easing matters when $(E,\mathcal{E})$ is uncountable.

Theorem 1.1 below gives a precise characterization of the construction of MPPs, but before explaining the intuition behind the construction and stating the theorem formally some notation is developed. Let $Z_n = (T_1,\ldots,T_n,Y_1,\ldots,Y_n)$, $z_n = (t_1,\ldots,t_n,y_1,\ldots,y_n)$ and $(z_n,t) = (t_1,\ldots,t_n,t,y_1,\ldots,y_n)$. Further, define $K_E^{(n)}$ to be the set of valid sequences of length n, that is

$$K_E^{(n)} = \big\{z_n \in \overline{\mathbb{R}}_+^n \times \overline{E}^n : 0 < t_1 \leq \ldots \leq t_n \leq \infty, t_k < t_{k+1} \text{ if } t_k < \infty \\ \text{and } y_k \neq \nabla \text{ if and only if } t_k < \infty\big\}$$

and

$$J_E^{(n)} = \big\{(z_n,t) : \ z_n \in K_E^{(n)} \text{ and } t_n \leq t \text{ with } t_n < t \text{ if } t_n < \infty\big\}.$$

In order to be able to discuss transition probabilities the obvious σ-algebras given by their respective coordinate projections, $\mathcal{K}_E^{(n)}$ and $\mathcal{J}_E^{(n)}$, are defined. The transition probabilities for the time points are denoted by $P^{(n)}$ for $n \in \mathbb{N}$, while the marks' transition probabilities are denoted by $\pi^{(n)}$ for $n \in \mathbb{N}_0$.

Having the formal setup in place allows us to discuss the construction of MPPs. The key ingredient is to begin with a marginal distribution $P^{(0)}$ of the first jump time, T_1, and then use transition probabilities to specify T_{n+1} given Z_n and Y_{n+1} given (Z_n,T_{n+1}). Note, that the distribution of the $n+1$th mark is conditioned on at what time the jump takes place T_{n+1} in accordance with the definition of $J_E^{(n)}$.

Theorem 1.1 (MPP Construction)**.** *Given a probability $P^{(0)}$ and a set of transition probabilities $P^{(n)}$, $n \geq 0$, and $\pi^{(n)}$, $n \geq 1$ satisfying*

$$\begin{aligned} P_{z_n}^{(n)}(]t_n, \infty]) &= 1 \qquad \text{if } t_n < \infty \\ P_{z_n}^{(n)}(\{\infty\}) &= 1 \qquad \text{if } t_n = \infty \\ \pi_{z_n}^{(n)}(E) &= 1 \qquad \text{if } t < \infty \\ \pi_{z_n}^{(n)}(\{\nabla\}) &= 1 \qquad \text{if } t = \infty \end{aligned} \tag{1.9}$$

there exists a unique probability $\overline{R}$ on $(\overline{K}_E, \overline{\mathcal{K}}_E)$ (potentially with explosions) such that $T_1^\circ(\overline{R}) = P^{(0)}$. Further, for $z_n \in K_E^{(n)}$ with $n \geq 1$, $P_{z_n}^{(n)}$ is a regular conditional distribution of T_{n+1}° given $Z_n^\circ = z_n$ and, similarly, for $(z_n, t) \in J_E^{(n)}$ with $n \geq 0$, $\pi_{z_n,t}^{(n)}$ is a regular conditional distribution of Y_{n+1}° given $(Z_n^\circ, T_{n+1}^\circ) = (z_n, t)$.

Finally, $\overline{R}$ defines a MPP without explosion if and only if

$$\overline{R}\left(\lim_{n\to\infty} T_n^\circ = \infty\right) = 0.$$

Proof. See [Ja06] Theorem 3.2.1 p. 22-23 for a discussion. □

Thus far, we have not spend a great deal of effort describing MPPs with explosion, but nevertheless Theorem 1.1 constructs an MPP that potentially has explosions. While a priori it is not clear at all whether the construction provided by the theorem actually exhaust all possible MPPs it is, unfortunately, not even possible to give a simple characterization of choices of $P^{(n)}$, $n \geq 0$, and $\pi^{(n)}$, $n \geq 1$, such that the resulting MPP is a genuine MPP without explosion. This is often referred to as the *stability problem* and an MPP (or RCM for that matter) without explosions is said to be *stable*.

It should also be noted that it is very deliberate that the theorem states "a" rather than "the" regular conditional distribution. Several choices of transition probabilities may result in the same $\overline{R}$ or R, respectively, if the differences are limited to a $\overline{R}$ or R null-set. Further, as (1.9) dictates the behavior of the transition probabilities when either $t_n = \infty$ or $t = \infty$, it suffices to describe the transition probabilities for $t_n, t < \infty$.

1.3. Continuous stochastic calculus

Our goal is to analyze jump-diffusions and while the above offers one possible description of jumps this sections deals with the diffusion parts needed. We first offer a basic understanding of the concepts needed for doing stochastic calculus inspired by [Ok00], while the more in-depth analysis of homogeneous one-dimensional diffusions and in particular the stationarity results are based on [Ja08]. Examples are condensed at the end in Section 1.3.4 where diffusions used later on are described.

As a motivating example, we begin by studying the solutions to the stochastic differential equation (SDE)

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \qquad X_0 = U, \tag{1.10}$$

where U is a random variable representing the boundary condition of the SDE and $\mathbf{B} = (B_t)_{t\geq 0}$ denotes the Brownian motion. The functions $b : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to]0, \infty[$ are called the drift and variance coefficient; in order to ensure existence

and uniqueness they are subject to restrictions (see Theorem 1.5). It is the properties of $\mathbf{X} = (X_t)_{t\geq 0}$ we shall study.

When the stochastic process described by (1.10) has well-defined expectation, we denote it by $\mathrm{E}_U[X_t]$. That is, the expectation of the stochastic variable X_t where $X_0 \overset{\mathcal{D}}{=} U$ as the expectation obviously depends on the boundary condition. If $U \equiv \varepsilon_x$ is degenerate we simple write E_x rather than $\mathrm{E}_{\varepsilon_x}$.

It is easily motivated why stochastic behavior should be introduced. Firstly, the phenomenon we wish to model may itself be subject to randomness. For example, when modeling the size of a large population there surely are some reasonable assumptions to be made about the growth of the population, but even for a large population some macro factors occurring at random such as natural disasters or lethal viral infections could influence the growth in the model.

In physics, it might be argued, phenomena exist that are very close to deterministic, but even a model of such could use the stochastic nature of the differential equation to model measurement uncertainties.

Yet another way to justify the introduction of a non-deterministic quantity in a model is by focusing on the major factors of interest considering the less significant ones in a compound variable of stochastic behavior. Take for instance the insulin producing beta cells in the pancreas. Certainly it seems reasonable to let the rate of insulin production depend on the glucose level of the blood while the sum of the myriad of other internal and external factors are treated stochastically.

In analogy to classical calculus, integration is imperative in attacking SDEs and our first task is therefore to give meaning to stochastic integrals. Imagine for a moment that (1.10) is an ODE. We would integrate on both sides and obtain

$$X_t = \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s,$$

where the entity $\int_0^t \sigma(s, X_s)dB_s$ is a stochastic integral. Both the integrand as well as the measure/increment function, integrator, depend on ω and as such the integral itself is a stochastic variable depending on t. Hence integrating one stochastic process $(X_t)_{t\geq 0}$ with respect to another stochastic process $(B_t)_{t\geq 0}$ returns a new stochastic process. It is the resulting stochastic process that, when well-defined, we wish to investigate.

Already at this point a potential problem lurks beneath the surface. If F is any non-decreasing non-negative function we know how to interpret $\int dF(x)$ as limit of a sum. Due to the assumptions on F we can choose any point in the interval considered as part of the limit and still end up with the same result. Any continuous function of finite variation can be written as a difference between two non-decreasing non-negative functions. The Brownian motion, however, has infinite variation and is nowhere monotone on any interval. Thus we might (and indeed will!) face a problem when deciding what point of an interval to use for our limit.

The different ways to define the stochastic integral turn out not to be equivalent and this thesis uses the Itô integral exclusively. Other prominent choices such as the Stratonovich integral exists for which the reader is referred to [Ok00].

1.3.1. Construction of the Itô integral

This section constructs the Itô integral with respect to a standard Brownian motion $\mathbf{B}$ in a rigorous manner. This suffices for the work presented in this thesis however, more general processes can be used as integrators treated in [Ja08] and [RW00].

We have already hinted that a nave approach imitating the construction of a standard integral could fail. Nevertheless, similarities are found. Specifically, a three-step construction is used where the first task is to replace the simple functions from Lebesgue integration with a different class of functions up to the task. These are called elementary functions. This is hardly surprising, but a crucial difference in the construction lies in the *Itô isometry*. It allows us to identify processes with finite quadratic variation with functions in $L^2(\mathbb{P})$, a space much more familiar to us.

Before proceeding with the isometry and the actual steps, we define the class of functions needed to build the Itô integral .

Definition 1.2. Define $\mathcal{V}(S,T)$ as the class of functions $f :]0,\infty[\times\Omega \to \mathbb{R}$ fulfilling:

- $f(t,\omega)$ is measurable for any $t \geq 0$ and $\omega \in \Omega$,
- $f(t,\omega)$ is $\mathcal{F}_t$-adapted, and
- $\mathrm{E}[\int_S^T f(t,\omega)^2 dt] < \infty$.

Note that the elements of $\mathcal{V}$ are random variables. By

$$\Pi_n = \{t_0 = S;\, t_1, \ldots, t_n = T\}, \qquad t_i < t_i + 1,$$

we define a partition of $[S,T]$ with mesh going to zero as n goes to infinity.

Definition 1.3. Let $\Pi_n(S,T)$ be a partition and Z_j a sequence of $\mathcal{F}_{t_j}$-adapted random variables. A function $\varphi \in \mathcal{V}(S,T)$ is called *elementary* if it can be written as

$$\varphi(t,\omega) = \sum_{j:t_j \in \Pi_n(S,T)} Z_j(\omega) 1_{[t_j, t_{j+1}[}(t).$$

Note that, if Z_j were not $\mathcal{F}_{t_j}$-adapted for all j, φ would not lie in $\mathcal{V}$. Recklessly, we define

$$\int_S^T \varphi(t,\omega)\, dB_t = \sum_{j:t_j \in \Pi_n(S,T)} Z_j(\omega)(B_{t_{j+1}} - B_{t_j})(\omega) \tag{1.11}$$

where it should be noted that the integrand is evaluated at the left endpoint. We need a special case of the Itô isometry for φ elementary and bounded, while a more general version is presented in Theorem 1.3.

Lemma 1.2 (Itô isometry, simple)**.** *Let φ be an elementary and bounded function in $\mathcal{V}(S,T)$. Then*

$$\mathrm{E}\left[\left(\int_S^T \varphi(t,\omega)\, dB_t\right)^2\right] = \mathrm{E}\left[\int_S^T \varphi^2(t,\omega)\, dt\right]. \tag{1.12}$$

Proof. Straight-forward calculations yield

$$\begin{aligned}\mathrm{E}\left[\left(\int_S^T \varphi(t,\omega)\, dB_t\right)^2\right] &= \mathrm{E}\left[\sum_{i,j} Z_i Z_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\right] \\ &= \sum_j \mathrm{E}[Z_i^2](t_{j+1} - t_j) = \mathrm{E}\left[\int_S^T \varphi^2(t,\omega)\, dt\right],\end{aligned}$$

where we have used that the standard Brownian motion has independent increments with mean 0. □

Note how the isometry takes us from unfamiliar domain—integration with respect to a stochastic process—to the well-acquainted space of $L^2(\mathbb{P})$. We now introduce the three-step process to approximate any function $f \in \mathcal{V}(S,T)$ by a sequence of bounded elementary functions. This work lays the basis for a rigorous definition of the Itô integral with respect to a Brownian motion.

Step 1 First, we wish to approximate $g \in \mathcal{V}(S,T)$ bounded and t-continuous by a sequence of elementary functions φ_n such that

$$\mathrm{E}\left[\int_S^T (g-\varphi_n)^2\,dt\right] \to 0 \qquad \text{for } n \to \infty$$

already from the beginning exploiting the convenience of operating on the familiar $L^2(\mathbb{P})$-space. Put $\varphi_n(t,\omega) = \sum_j g(t_j,\omega)1_{[t_j,t_{j+1}[}(t)$ with $t_j = t_j^n \in \Pi_n$. From continuity of g it follows that $\int_S^T (g-\varphi_n)^2\,dt \to 0$ for $n \to \infty$ for all ω by dominated convergence.

Step 2 Let $h \in \mathcal{V}(S,T)$ be bounded. Then there exists a sequence of g_n bounded and t-continuous from Step 1 such that

$$\mathrm{E}\left[\int_S^T (h-g_n)^2\,dt\right] \to 0 \qquad \text{for } n \to \infty.$$

Begin by defining the auxiliary functions $\psi_n(t)$ continuous with support on $(-1/n, 0)$ and integrating to 1. Let

$$g_n(t,\omega) = \int_0^t \psi_n(s-t)h(s,\omega)\,ds.$$

It is immediate that if $|h| \le M$ then $|g_n| \le M$, too. Realizing that g_n is t-continuous is simple. For $t_k \to t$ we find that $1_{0,t_k}(s)h(s) \to 1_{0,t}(s)h(s)$. The latter serving as an integrable upper bound, dominated convergence implies that $g_n(t_k,\omega) \to g_n(t,\omega)$ for each ω. It is also needed that g_t is $\mathcal{F}_t$-adapted in order for g_t to be in $\mathcal{V}(S,T)$. This follows from $h \in \mathcal{V}(S,T)$, is a bit technical, and omitted here. Using dominated convergence once more it is seen that $\int_S^T (h-g_n)^2\,dt \to 0$ for $n \to \infty$ for each ω.

Step 3 Finally, let $f \in \mathcal{V}(S,T)$ with no further restriction. It is then a simple matter to construct a sequence h_n each in $\mathcal{V}(S,T)$ and bounded such that

$$\mathrm{E}\left[\int_S^T (f-h_n)^2\,dt\right] \to 0 \qquad \text{for } n \to \infty.$$

Define h_n as follows and apply dominated convergence

$$h_n(t,\omega) = \begin{cases} -n & f(t,\omega) < -n, \\ f(t,\omega) & \text{otherwise}, \\ n & f(t,\omega) > n. \end{cases}$$

This concludes the steps needed to construct the Itô integral. Using Step 1 to 3, we know there exists a sequence of elementary functions φ_n such that $\int_S^T (f-\varphi_n)^2\,dt \to 0$.

Definition 1.4 (Itô integral)**.** Let $f \in \mathcal{V}(0,t)$ and $\mathbf{B}$ be a Brownian motion. The Itô integral of f with respect to $\mathbf{B}$ is defined as

$$(f \cdot \mathbf{B})_t = \int_0^t f(s,\omega)\, dB_s = \lim_{n\to\infty} \int_0^t \varphi_n(s,\omega)\, dB_s \tag{1.13}$$

We know how to interpret the right-hand side of (1.13) and that it can be identified with an element of $L^2(P)$ using (1.12). Specifically, it is a Cauchy sequence in $L^2(P)$ from which it follows that the limit exists and is well-defined, that is independent of the choice of φ_n's. From the isometry for elementary functions (1.12), it follows immediately that

Theorem 1.3 (Itô isometry)**.** *For any $f \in \mathcal{V}(S,T)$ it holds that*

$$\mathrm{E}\left[\left(\int_S^T f(t,\omega)\, dB_t\right)^2\right] = \mathrm{E}\left[\int_S^T f^2(t,\omega)dt\right]. \tag{1.14}$$

Note, that unlike in classical calculus it matters greatly at what time point we choose to evaluate the integrand. The Itô integral given in (1.11) uses the left-most point of an interval $[t_i, t_{i+1}[$ that is t_i, whereas the Stratonovich integral uses the midpoint $(t_i + t_{i+1})/2$. When discussing existence of a (stochastic) integral there are two parts to be considered: The integrand and the integration measure/process. As mentioned, it is possible to integrate with respect to other processes than the Brownian motion. The study of when (1.13) converges is not trivial at all and omitted.

1.3.2. Properties of the Itô integral

Unsurprisingly, from the definition of elementary stochastic integrals (1.11), we observe that they are linear and that additivity of integrals holds. Perhaps more interestingly, it is seen that the integral is $\mathcal{F}_T$-measurable and a back-of-the-envelope calculation shows that for elementary functions

$$\mathrm{E}\left[\int_S^T \varphi(t,\omega)\, dB_t\right] = \sum_j \mathrm{E}[Z_j]\,\mathrm{E}[B_{t_{j+1}} - B_{t_j}] = 0$$

as $\mathbf{Z} \perp \mathbf{B}$. By taking limits these properties carries over to the Itô integral for all functions in $\mathcal{V}(S,T)$.

Before discussing stochastic integration in details, we need another definition. In real analysis integrating a polynomial again returns a polynomial, that is, the Riemann-Stieltjes integral operator maps from the class of polynomials into itself. In the same way, it is convenient to describe a class of processes, which are mapped into itself using the Itô integral operator.

Definition 1.5 (Itô Process, 1 dimensional)**.** A stochastic process $\mathbf{X}$ is called an Itô process if it, in differential form, can be written as

$$dX_t = f(t,\omega)dt + g(t,\omega)dB_t, \tag{1.15}$$

where B_t is a 1-dimensional Brownian motion and f and g are $\mathcal{F}_t$-adapted. Equivalently, in integral form

$$X_t = X_0 + \int_0^t f(s,\omega)ds + \int_0^t g(s,\omega)dB_s.$$

Before stating the infamous Itô's formula we need to recall yet another definition.

Definition 1.6 (Quadratic (co)variation)**.** Suppose $\mathbf{X}$ and $\mathbf{Y}$ are real-valued stochastic processes then the *quadratic covariation* on the interval $[s,t]$ is defined as, when the limit exists,

$$[\mathbf{X},\mathbf{Y}]_{[s,t]} := \lim_{|\Pi|\to 0} \sum_{t_i,t_{i+1}\in\Pi} (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i}), \tag{1.16}$$

where Π is a partition of $[s,t]$, $|\Pi|$ denotes the mesh length, and convergence is in probability. Define further the *quadratic variation* as

$$[\mathbf{X}]_{[s,t]} := [\mathbf{X},\mathbf{X}]_{[s,t]}. \tag{1.17}$$

Note that the quadratic (co)variation is a stochastic variable.

The limit always exists for Itô processes and with X_t given by (1.15) one can show that

$$d[\mathbf{X}]_t = g^2 dt \quad \text{and in particular} \quad d[\mathbf{B}]_t = dt. \tag{1.18}$$

While Definition 1.4 is rarely used for direct computations of integrals, Itô formula below is. It also shows that Itô process are stable under transformation by (sufficiently smooth) functions described in the following theorem:

Theorem 1.4 (Itô's formula, 1 dimensional)**.** *Let $\mathbf{X}$ be an Itô process and $h(t,x)$ be once differentiable in t and twice in x. With $Y_t = h(t,X_t)$, $\mathbf{Y}$ is again an Itô process given in differential form by*

$$dY_t = \frac{dh}{dt}(t,X_t)dt + \frac{dh}{dx}(t,X_t)dX_t + \frac{1}{2}\frac{d^2h}{d^2x}(t,X_t)d[\mathbf{X}]_t, \tag{1.19}$$

where $[\mathbf{X}]_t$ denotes the quadratic variation found in (1.18).

Proof. [Ok00] p. 44-48 offers a proof sketch for the simplified setting considered here. For a rigorous proof see [RW00] p. 60-62 treating a more general case. □

When first having defined the Riemann-integral, the urge to calculate $\int_0^t x\,dx$ is overwhelming. In analogy, we wish to calculate $\int_0^t B_s\,dB_s$. Inspired by the classical situation, we apply Itô's formula to $g(t,x) = \frac{1}{2}x^2$ such that $Y_t = \frac{1}{2}B_t^2$. Plugging into (1.19), we obtain

$$d\frac{1}{2}B_t^2 = 0 + B_t dB_t + \frac{1}{2}d[\mathbf{B}]_t = B_t dB_t + \frac{1}{2}dt,$$

which is equivalent to

$$\int_0^t B_s dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}\int_0^t dx = \frac{1}{2}B_t^2 - \frac{1}{2}t. \tag{1.20}$$

The latter term shows that the Itô integral behaves differently from the ordinary Riemann-Stieltjes integral. It is also an example of a process of the form $dX_t = g dB_t$ where the integrated process cannot be written on the same form.

Finally, note that not all computations—albeit choosing g cleverly—leads to expressions as simple as (1.20). For example with $g(t,x) = x^3/3$, Itô formula yields

$$\int_0^t B_s^2 dB_s = \frac{1}{3}B_t^3 - \int_0^t B_s ds.$$

Nevertheless, it often pays to choose g along the lines of the corresponding anti-derivative from classical calculus.

To conclude the section an existence and uniqueness theorem is offered.

Theorem 1.5. *Let $b(t,x)$ and $\sigma(t,x)$ both be Lipschitz continuous, that is*

$$|b(t,x) - b(t,y)| \leq K_b|x-y|,$$
$$|\sigma(t,x) - \sigma(t,y)| \leq K_\sigma|x-y|$$

for $t \geq 0$ and $x, y \in \mathbb{R}$ with K_b and K_σ constants. Then (1.10)

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dB_t, \qquad X_0 = U$$

has a unique solution for each boundary condition given by a $\mathcal{F}_0$-measurable random variable U.

Proof. See [Ja08] p. 75-81. □

The proof is lengthy and the requirements for b and σ are also stronger than needed; [Ja08] p. 81-82 discusses details.

1.3.3. Analysis of one-dimensional homogeneous diffusions

This section in brevity establishes properties of diffusion processes and is based on Chapter 11 of [Ja08] where the reader may find most proofs and further details. A diffusion is a process solving a particular kind of SDE of the form

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \qquad X_0 = x_0 \tag{1.21}$$

with B a standard Brownian motion. Note that b and σ are no longer allowed to depend on t, but solely X_t. We do not assume b and σ to be Lipschitz and for instance the Cox-Ingersoll-Ross process (see Section 1.3.4.4) fails to meet this criterion. It is assumed that the diffusion lives on an interval $]l,r[$.

Before proceeding, we define the differential operator A for $f \in C^2(\mathbb{R})$

$$Af(x) = b(x)f'(x) + \frac{\sigma^2(x)}{2}f''(x). \tag{1.22}$$

Note that, when $h(t,x) = h(x) \in C^2(\mathbb{R})$ does not depend on t, Itô's formula to gives

$$dh(X_t) = Ah(X_t)dt + h'(X_t)\sigma(X_t)dB_t, \tag{1.23}$$

as the dh/dt-term is zero.

The process $\mathbf{X}$ being the one of interest, it turns out that considering the transformed process $S(\mathbf{X})$ for certain S yields results about $\mathbf{X}$ itself via Itô's formula. Let therefore $S \in C^2(]l,r[)$ where we assume that $X_t \in]l,r[$ for all t. (1.23) becomes

$$dS(X_t) = AS(X_t)dt + S'(X_t)\sigma(X_t)dB_t.$$

A sufficient condition for $S(\mathbf{X})$ to be a continuous local martingale is that $AS \equiv 0$, equivalent to a simple first-order homogeneous ordinary differential equation in S' with solution

$$S'(x) = c\exp\left(-\int_{x_0}^{x}\frac{2b(y)}{\sigma^2(y)}dy\right) \tag{1.24}$$

for some c and unique up to an affine transformation. A function S fulfilling (1.24) is called a *scale function*. Usually $c > 0$ is chosen such that S is increasing. In particular this implies that both

$$S(l) := \lim_{y \downarrow l}(y) \geq -\infty \quad \text{and} \quad S(r) := \lim_{y \uparrow r}(y) \leq \infty \tag{1.25}$$

are well-defined.

In general, for a stopping time τ and a process $\mathbf{X}$ define the *stopped process* $\mathbf{X}^\tau$ by $X_t^\tau := X_{\tau \wedge t}$. Put

$$\begin{aligned} \tau_{a,b} &:= \inf\{t : X_t = a \text{ or } X_t = b\}, \\ \tau_a &:= \inf\{t : X_t = a\}, \qquad\qquad l < a < x_0 < b < r \end{aligned}$$

with the convention that $\inf\{\emptyset\} = \infty$. It is clear that $S(\mathbf{X})^{\tau_{a,b}}$ is a bounded local martingale and thus a real martingale which implies

$$\mathrm{E}_{x_0}[S(X_{t \wedge \tau_{a,b}})] = \mathrm{E}(S(X_0)) = S(x_0)$$

for all t. Using dominated convergence for $t \to \infty$ we find that $\mathrm{E}_{x_0}[S(X_{\tau_{a,b}})] = S(x_0)$. To connect back to $\mathbf{X}$ observe that ($\lim_{t\to\infty} S(X_t)$ is well-defined by the martingale convergence theorem)

$$S(X_{\tau_{a,b}}) = \begin{cases} S(b) & \text{on } \tau_b < \tau_a \\ S(a) & \text{on } \tau_a < \tau_b \\ \lim_{t\to\infty} S(X_t) & \text{on } \tau_{a,b} = \infty. \end{cases} \tag{1.26}$$

and as [Ja08] p. 114-115 shows that $\mathbb{P}_{x_0}(\tau_{a,b} = \infty) = 0$ for all $x_0 \in]l, r[$, it holds that

$$S(x_0) = \mathrm{E}_{x_0}[S(X_{\tau_{a,b}})] = S(b)\mathbb{P}_{x_0}(\tau_b < \tau_a) + S(a)\mathbb{P}_{x_0}(\tau_a < \tau_b).$$

Using that $(\tau_b < \tau_a)$ is the complement of $(\tau_a < \tau_b)$ with probability 1 and rearranging gives

$$\mathbb{P}_{x_0}(\tau_b < \tau_a) = \frac{S(x_0) - S(a)}{S(b) - S(a)}. \tag{1.27}$$

Note first, that this expression is well-defined as it yields the same result even if S is transformed affinely. Secondly, it is a result concerning the original process $\mathbf{X}$ reaching the values a or b. If $S(x) = x$ the corresponding diffusion is said to be in natural scale and (1.27) becomes particular simple. [RW00] p. 275-276 proof that $S(\mathbf{X})$ is always in natural scale.

Closely related to the scale function is the speed measure defined as

$$k(x) = \frac{2}{\sigma^2(x)S'(x)}. \tag{1.28}$$

If the scale function is affinely transformed to $c_1 + c_2 S$, k has to be replaced by k/c_2. For a standard Brownian motion it is easily verified that (see Section 1.3.4.1)

$$S'(x) \equiv 1 \quad \text{implying} \quad S(x) = x \quad \text{and} \quad k(x) = 2.$$

This example helps grasp the results quoted in the following theorem, which is stated using the so-called *Green functions* symmetric in x and y and given by

$$G_{a,b}(x, y) = \frac{[S(x) - S(a)][S(b) - S(y)]}{S(b) - S(a)} \qquad \text{for } a \leq x \leq y \leq b.$$

Theorem 1.6. *Let* $\mathbf{X}$ *be a diffusion as in (1.21) with values in* $]l, r[$ *and put* $x_0 = X_0$. *For* $l < a < x < b < r$ *it holds that* $\mathbb{P}_{x_0}(\tau_{a,b} = \infty) = 0$ *and*

$$\mathbb{P}_{x_0}(\tau_b < \tau_a) = 1 - \mathbb{P}_{x_0}(\tau_a < \tau < b) = \frac{S(x_0) - S(a)}{S(b) - S(a)}.$$

Further, for $\varphi : [a, b] \to \mathbb{R}$ *bounded and measurable*

$$\mathrm{E}_{x_0}\left[\int_0^{\tau_{a,b}} \varphi(X_s) ds\right] = \int_a^b G_{a,b}\, \varphi(y) k(y)\, dy, \tag{1.29}$$

and in particular (choosing $\varphi \equiv 1$*)*

$$\mathrm{E}_{x_0}\left[\tau_{a,b}\right] = \int_a^b G_{a,b}\, k(y)\, dy \tag{1.30}$$

Proof. See [Ja08] p. 115-116. □

Note that G, k and S all solely depend on b and σ. The additional notation is simply a way to make (1.29) and (1.30) more readable. Applying Theorem 1.6 to the standard Brownian motion, we find that

$$\mathbb{P}_{x_0}(\tau_b < \tau_a) = \frac{x_0 - a}{b - a} \quad \text{and in particular} \quad \mathbb{P}_0(\tau_1 < \tau_{-1}) = \frac{1}{2}$$

unsurprisingly stating that a standard Brownian motion is equally likely to cross -1 or 1 first. Further,

$$\mathrm{E}_{x_0}[\tau_{a,b}] = (x_0 - a)(b - x_0) \quad \text{in particular} \quad \mathrm{E}_0[\tau_{-1,1}] = 1,$$

which interpretation is that the expected time for a standard Brownian motion to hit either -1 or 1 is 1.

Recalling that we no longer have assumed b and σ to be Lipschitz we are still in deep water concerning the existence and uniqueness of a diffusion on $]l, r[$. The context of the following theorem is a bit unclear since we have omitted a significant part of the theory leading up to it. Nevertheless, it is easy to state and check. Without further ado and omitting the proof:

Theorem 1.7. *Fix* $]l, r[$ *and assume that continuous* $b :]l, r[\to \mathbb{R}$ *and* $\sigma :]l, r[\to]0, \infty[$ *are such that*

$$S(l) = -\infty \quad \text{or} \quad \int_l^y [S(z) - S(l)] k(z)\, dz = \infty, \qquad y \in]l, r[\tag{1.31}$$

$$S(r) = \infty \quad \text{or} \quad \int_y^r [S(r) - S(z)] k(z)\, dz = \infty, \qquad y \in]l, r[\tag{1.32}$$

If $\mathbf{B}$ *is a standard Brownian motion on* $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ *and* U *is* $\mathcal{F}_0$*-measurable with values in* $]l, r[$ *then*

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t, \qquad X_0 \equiv U$$

has a unique solution.

Proof. The result comes via the construction of regular diffusions using time changes as done in [IM65] section 5.2, p. 167-171. □

This not only gives a taste of the fundamental use of the scale function and speed measure, but also a very explicit condition to check for a given diffusion. Recall that S and k solely depend on b and σ. The idea of Theorem 1.7 is that if one succeeds in producing two functions b and σ along with two endpoint l and r that satisfy both (1.31) and (1.32) then we are certain that a diffusion uniquely exists if only we start it according to a U attaining values in $]l, r[$.

We conclude the section with the following result of significant importance.

Theorem 1.8. *The diffusion* $\mathbf{X}$ *has stationary initial distribution* ν *if and only if*

$$K = \int_r^l k(x)dx < \infty \tag{1.33}$$

in which case

$$\nu(dx) = \frac{1}{K}k(x)dx. \tag{1.34}$$

In particular it is necessary that $\mathbf{X}$ *is recurrent, that is each level between* $]l, r[$ *is attained infinitely often in* $[t, \infty[$ *for any* t*.*

Proof. A partial proof is found in [Ja08] theorem 11.7 p. 123-124. □

1.3.4. Diffusion examples

The most famous example of a diffusion is likely the Brownian motion, but there are of course other or interest and some of these are used later on. Here we apply the just-developed theory to obtain some properties. Many useful results are compiled in [BS02].

1.3.4.1. Brownian motion

We very briefly consider the Brownian motion itself as no exposition of stochastic calculus would be complete without it. It of course solves the SDE

$$dX_t = dB_t, \qquad X_0 = 0.$$

We already know that $S(x) = x$ and $k(x) = 2$. From Theorem 1.8 it is seen that no stationary distribution for the Brownian motion exists.

A general Brownian motion solves with drift θ and variance σ

$$dX_t = \mu X_t + \sigma dB_t$$

with scale function

$$S(x) = \begin{cases} x & \text{if } \theta = 0, \\ -\frac{\sigma^2}{2\theta}e^{-2\theta/\sigma^2 x} & \text{otherwise} \end{cases} \tag{1.35}$$

and speed measure

$$k(x) = \frac{2}{\sigma^2}e^{2\theta/\sigma^2 x}. \tag{1.36}$$

The transition probabilities (see (1.55) are of particular interest to us and given by

$$p_t(x, y) = \frac{1}{\sqrt{2\pi\sigma^2 t}}e^{-\frac{(x-y)^2}{2\sigma^2}}. \tag{1.37}$$

1.3.4.2. Geometric Brownian motion

The geometric Brownian motion (GBM) is described by

$$dX_t = \theta X_t dt + a X_t dB_t, \qquad X_0 = x_0 > 0 \tag{1.38}$$

with $\theta, a \in \mathbb{R}$. Our first claim is that it is always positive that is $l = 0$ and $r = \infty$. In order to show this we need the scale function and speed measure. Using formula (1.24) we find for $x > 0$

$$S'(x) = \exp\left(-\frac{2\theta}{a^2}\log x\right) = x^{-2\theta/a^2}$$

and

$$S(x) = \begin{cases} \frac{a^2}{a^2-2\theta}x^{1-2\theta/a^2} & 2\theta \neq a^2, \\ \log(x) & 2\theta = a^2. \end{cases}$$

Hence the speed measure (1.28) becomes

$$k(x) = \frac{2}{a^2}x^{2\theta/a^2-2}.$$

We look for conditions on θ and a such that (1.31) and (1.32) are fulfilled in order to use Theorem 1.7. To calculate $S(0)$ recall from (1.25) the convention that $S(0) := \lim_{y\downarrow 0} S(y)$. For $2\theta \neq a^2$ it is seen that $S(0) = 0$ and for all $y \in]0, \infty[$

$$\int_0^y [S(z) - S(0)]k(z)\,dz = \int_0^y S(z)k(z)dz = \int_0^y Cz^{-1}\,dz = \infty$$

for some $C > 0$. If $2\theta = a^2$ then $S(0) = \infty$. The left border $l = 0$ therefore yields no further restrictions.

For $r = \infty$, we see that $S(\infty) = \infty$ if and only if $2\theta \leq a^2$ and otherwise $S(\infty) = 0$. The diffusion therefore stays on $]0, \infty[$ if $2\theta \leq a^2$, but what about $2\theta > a^2$? Let this be the case and observe that $S(x) < 0$ for $x > 0$. Therefore

$$\int_y^\infty [S(\infty) - S(z)]k(z)\,dz = \int_y^\infty -S(z)k(z)dz = \int_y^\infty C'\frac{1}{z}\,dz = \infty$$

with $C' > 0$. We conclude that the GBM in (1.38) is positive for all choices of θ and a.

Hence we can rest-assured transform X_t with $g(t, x) = \log x$ using Itô's formula, which gives

$$d(\log X_t) = \frac{dX_t}{X_t} - \frac{1}{2}\frac{1}{X_t^2}d[\mathbf{X}]_t = \frac{dX_t}{X_t} - \frac{a^2}{2}dt = \theta dt + a dB_t - \frac{a^2}{2}dt$$

as $d[\mathbf{X}]_t = a^2 X_t^2 dt$. This differential form is equivalent to

$$[\log X_s]_0^t = (\theta - \frac{a^2}{2})t + aB_t.$$

Rearranging and taking the exponential we obtain the solution

$$X_t = X_0 \exp\left[(\theta - \frac{a^2}{2})t + aB_t\right], \tag{1.39}$$

from which it is clear that the GBM is strictly positive.

Hence its expectation exists and is

$$\mathrm{E}[X_t] = X_0 e^{\theta t}. \tag{1.40}$$

To conclude the analysis of the GBM we quote [BS02] p. 606 for X_T where $T \sim \exp\lambda$ is independent of $\mathbf{X}$. Thus we consider the GBM with exponential stopping and its density is given by

$$\mathbb{P}_x(X_T \in dz) = \begin{cases} \frac{\lambda}{z\sigma^2\sqrt{\eta^2+2\lambda/\sigma^2}} \left(\frac{x}{z}\right)^{\sqrt{\eta^2+2\lambda/\sigma^2}-\eta} dz & x \le z \\ \frac{\lambda}{z\sigma^2\sqrt{\eta^2+2\lambda/\sigma^2}} \left(\frac{z}{x}\right)^{\sqrt{\eta^2+2\lambda/\sigma^2}+\eta} dz & z \le x. \end{cases} \tag{1.41}$$

with $\eta = \frac{\theta}{\sigma^2} - \frac{1}{2}$.

1.3.4.3. Ornstein-Uhlenbeck process

The Ornstein-Uhlenbeck process (OU) is governed by the SDE

$$dX_t = (a - bX_t)dt + \sigma dB_t, \qquad X_0 = x_0 \tag{1.42}$$

with $a \in \mathbb{R}$, and $b, \sigma > 0$.

The OU process can be made stationary . The stationary distribution is normal with mean a/b and variance $\frac{\sigma^2}{2b}$; in particular the stationary density is

$$u(x) = \frac{1}{\sqrt{\pi\frac{\sigma^2}{b}}} \cdot \exp\left(\frac{(x-a/b)^2}{\frac{\sigma^2}{b}}\right). \tag{1.43}$$

The transition probabilities of the OU process are given by (see [BS02] p. 137)

$$(X_{t+s}|X_s = x) \overset{\mathcal{D}}{=} \mathcal{N}\left(xe^{-bt} + \frac{a}{b}(1-e^{-bt}), \frac{\sigma^2}{2b}(1-e^{-2bt})\right) \tag{1.44}$$

such that their densities are

$$p_t(x, dy) = \frac{1}{\sqrt{\pi\frac{\sigma^2}{b}(1-e^{-2bt})}} \exp\left(-\frac{\left[y - xe^{-bt} + \frac{a}{b}(1-e^{-bt})\right]^2}{\frac{\sigma^2}{b}(1-e^{-2bt})}\right) dy. \tag{1.45}$$

In particular

$$\mathrm{E}_{x_0}[X_t] = x_0 e^{-bt} + \frac{a}{b}(1-e^{-bt}). \tag{1.46}$$

from which it is seen that the OU process is mean reverting with asymptotic expectation a/b.

Using the parametrization of the OU-process $dX_t = \theta(-X_t dt + 2\sigma dB_t)$ for notational convenience and again quoting [BS02] p. 522 the OU-process at exponential stopping has density

$$\begin{aligned} \mathbb{P}_x(X_T \in dz) = {} & \frac{\lambda\Gamma(\lambda/\theta)}{\theta\sigma\sqrt{2\pi}} e^{(x^2-z^2)/4\sigma^2} \\ & \cdot D_{-\lambda/\theta}\left(-\frac{z+x-|z-x|}{2\sigma}\right) D_{-\lambda/\theta}\left(\frac{z+x+|z-x|}{2\sigma}\right) dz \end{aligned} \tag{1.47}$$

where $D_{-\eta}(x)$ is the parabolic cylinder function given by ([BS02] p. 639)

$$D_{-\eta}(x) = e^{-x^2/4}2^{-\eta/2}\sqrt{\pi}\Big[\frac{1}{\Gamma((\eta+1)/2)}\left(1+\sum_{k=1}^{\infty}\frac{\eta(\eta+2)\cdots(\eta+2k-1)}{(2k)!}x^{2k}\right) - \frac{x\sqrt{2}}{\Gamma(\eta/2)}\left(1+\sum_{k=1}^{\infty}\frac{(v+1)(v+3)\cdots(v+2k-1)}{(2k+1)!}x^{2k}\right)\Big].$$

1.3.4.4. Cox-Ingersoll-Ross process

The SDE

$$dX_t = (a+bX_t)dt + \sigma\sqrt{X_t}dB_t, \qquad X_0 = x_0 > 0 \tag{1.48}$$

with $a, b \in \mathbb{R}$ and $\sigma > 0$ describes the Cox-Ingersoll-Ross process (CIR).[1] We are interested in determining parameter sets such that the process stays strictly positive and is recurrent. It can be shown that they are $2a/\sigma^2 \geq 1$, $b < 0$ or $2a/\sigma^2 = 1$, $b = 0$; a fact we merely use here.

We first want to find the expectation value of the CIR-process, which exists as it is a positive process. We wish to show that $\sigma\,\mathrm{E}_{x_0}[\int_0^t \sqrt{X_s}dB_s] = 0$, true by the Itô isometry if $\mathrm{E}_{x_0}[\int_0^t \sqrt{X_s}^2 ds] < \infty$. To show this, define the stopping time $\tau_n := \inf\{t : X_t = n\}$. Later on we let $n \to \infty$ and therefore assume without loss of generality that $x_0 < n$. Thus

$$X_{\tau_n\wedge t} = x_0 + \int_0^{\tau_n\wedge t} a + bX_s\,ds + \sigma\int_0^t \sqrt{X_s}1_{(\tau_n\geq s)}\,dB_s.$$

A stochastic integral is always a local Martingale and using the stopping time τ_n it becomes a true Martingale. Therefore

$$\mathrm{E}_{x_0}[X_{\tau_n\wedge t}] = x_0 + \mathrm{E}_{x_0}\left[\int_0^{\tau_n\wedge t} a + bX_s\,ds\right] = x_0 + a\,\mathrm{E}_{x_0}[\tau_n\wedge t] + b\,\mathrm{E}_{x_0}\left[\int_0^{\tau_n\wedge t} X_s\,ds\right].$$

The case with $b = 0$ offers no additional information and we assume that $b < 0$. Letting $n \to \infty$ we see that $\mathrm{E}_{x_0}[\tau_n \wedge t] \to t$ and $\mathrm{E}_{x_0}[\int_0^{\tau\wedge t} X_s\,ds] \to \mathrm{E}_{x_0}[\int_0^t X_s ds]$, which is the entity we want to show is finite. The left-hand side is positive as the process is positive. Thus $\mathrm{E}_{x_0}[\int_0^t X_s\,ds] = \infty$ leads to a contradiction as $b < 0$ and the right-hand side would therefore be $-\infty$. Hence $\mathrm{E}_{x_0}[\int_0^t X_s\,ds] < \infty$ and we conclude that $\mathrm{E}_{x_0}[\int_0^t \sqrt{X_s}\,dB_s] = 0$. Knowing this we get

$$\mathrm{E}_{x_0}[X_t] = \mathrm{E}[x_0] + \int_0^t a - b\,\mathrm{E}_{x_0}[X_s]ds$$

implying

$$\frac{d}{dt}\mathrm{E}[X_t] = a - b\,\mathrm{E}[X_t], \qquad X_0 = x_0.$$

This is an ODE with solution

$$\mathrm{E}_{x_0}[X_t] = x_0e^{bt} - \frac{a}{b}(1-e^{bt}). \tag{1.49}$$

The CIR-process has asymptotic expectation a/b if $b < 0$.

[1] See [Ja08] Example 11.8 for details

Using Theorem 1.8 one finds that the CIR-process is stationary if and only if

$$\frac{2a}{\sigma^2} \geq 1 \quad \text{and} \quad b < 0 \tag{1.50}$$

with stationary density

$$u(x) = \frac{(-2b/\sigma^2)^{2a/\sigma^2}}{\Gamma(2a/\sigma^2)} x^{2a/\sigma^2 - 1} \exp\left(\frac{2b}{\sigma^2} x\right), \quad x > 0. \tag{1.51}$$

This is recognized as a gamma distribution with shape parameter $\alpha = 2a/\sigma^2$ and rate parameter $\beta = -2b/\sigma^2$.

1.4. Markov processes; stationarity and the infinitesimal generator

A Markov process in continuous time is a stochastic process $\mathbf{X} = (X_t)_{t \geq 0}$ with the Markov property. Loosely stated this means that the future only depends on the present or at every t a Markov process starts a fresh. We define the transition probability at time t denoted by P_t as

$$P_t f(X_s) := \mathrm{E}[f(X_t)|X_s] \tag{1.52}$$

with densities

$$p_{st}(x, dy) := \mathbb{P}(X_t \in dy | X_s = x). \tag{1.53}$$

Definition 1.7. A process $\mathbf{X}$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ is *Markovian* if for all t it holds for any $0 \leq s \leq t$ that there exists a transition probability p_{st} on G such that

$$\mathbb{P}(X_t \in C | \mathcal{F}_s) = p_{st}(X_s, C), \qquad G \in \mathcal{G}. \tag{1.54}$$

It is further called *time-homogeneous* if it only depends on $t - s$ in which case we are tempted to write

$$p_t(x, dy) := \mathbb{P}(X_{t+s} \in dy | X_s = x). \tag{1.55}$$

A fundamental result is the Chapman-Kolmogorov equation stating that

$$P_{t+s} = P_t P_s = P_s P_t, \quad s, t \geq 0$$

from which it is seen that the P_t's make up a semi group.[2]

Definition 1.8 (Stationarity)**.** A continuous time process $\mathbf{X}$ is called stationary if we for any $n \in \mathbb{N}$ and for any selection of n time points $0 \leq t_1 < \ldots < t_n$ and $s > 0$ have

$$(X_{t_1}, \ldots, X_{t_n}) \stackrel{\mathcal{D}}{=} (X_{t_1+s}, \ldots, X_{t_n+s}). \tag{1.56}$$

Two examples of stationary processes at each end of the spectrum are $\mathbf{Y} = (Y_t)_{t \geq 0}$ where all Y_t are iid and $\mathbf{Z} = (Z_t)_{t \geq 0}$ with $Z_{s_1} = Z_{s_2}$ for all s_1 and s_2.

Of particular interest in this thesis is the study of stationarity for Markov processes. For such a process the above definition is equivalent to X_t having distribution ρ for all t if the initial distribution, that is the distribution of X_0, is ρ. We then call ρ the *stationary distribution* or the *invariant distribution*.

[2] A semi group differs from a group in that for each element there is no guarantee that an inverse exists.

Definition 1.9 (Infinitesimal generator)**.** Let the *infinitesimal generator* A of a Markov process $\mathbf{X}$ be defined as the operator

$$Af(x) = \lim_{t \downarrow 0} \frac{\mathrm{E}_x[f(X_t)] - f(x)}{t} \tag{1.57}$$

for suitable f. Let $\mathcal{D}(A)(x)$ be the set of functions for which the limit exists at x, while $\mathcal{D}(A) = \cap_x \mathcal{D}(A)(x)$ is the set of functions where the limit exists for all x.

Further, we shall define a *determining class* as a class $\mathcal{D}$ of measurable and bounded functions such that if two probability measures ρ_1 and ρ_2 satisfies $\rho_1(h) = \rho_2(h)$ for all $h \in \mathcal{D}$ then $\rho_1 = \rho_2$. It is clear that h needs to be measurable. If h was not required to be bounded then $\rho_1(h)$ and $\rho_2(h)$ might both attain the value ∞ despite $\rho_1 \neq \rho_2$. It is well-known that $C_0^\infty(\mathbb{R})$ is dense in $L^2(\mathbb{R})$ (see Theorem A.2) and hence $C_0^\infty(\mathbb{R})$ is a determining class for probability measures on $L^2(\mathbb{R})$.

While (1.57) seems inaccessible at first, studying the particular case of an Itô diffusion sheds some light onto matters. Recall that the integration form of a diffusion is given by

$$X_t = X_0 + \int_0^t b(X_s)\, dx + \int_0^t \sigma(X_s)\, dB_s$$

where the Itô formula for $f \in C^2(\mathbb{R})$ gives

$$f(X_t) = f(X_0) + \int_0^t b(X_s) f'(X_s) + \frac{1}{2}\sigma^2(X_s) f''(X_s)\, ds + \int_0^t f'(X_s)\sigma(X_s)\, dB_s.$$

Consider $f(\mathbf{X})$ and condition on $X_0 = x$. Take expectation to get

$$\mathrm{E}_x[f(X_t)] = f(x) + \mathrm{E}_x\left[\int_0^t b(X_s) f'(X_s) + \frac{1}{2}\sigma^2(X_s) f''(X_s)\, ds\right]$$

recalling that the expectation of a stochastic integral with respect to the Brownian motion is 0 under suitable conditions. Inserting this into (1.57) yields

$$\begin{aligned} Af(x) &= \lim_{t \downarrow 0} \frac{1}{t} \int_0^t \mathrm{E}_x\left[b(X_s) f'(X_s) + \frac{1}{2}\sigma^2(X_s) f''(X_s)\right] ds \\ &= \mathrm{E}_x\left[b(X_0) f'(X_0) + \frac{1}{2}\sigma^2(X_0) f''(X_0)\right] \\ &= b(x) f'(x) + \frac{1}{2}\sigma^2(x) f''(x). \end{aligned}$$

This is very comforting! In (1.22) we introduced the differential operator A on $f \in C^2(\mathbb{R})$ given by $Af = bf' + \frac{1}{2}\sigma^2 f''$ and the above shows that for such f the two coincide.

Interestingly, it turns out that a sufficient and necessary condition for ν to be an invariant distribution is

$$\nu(Af) = 0, \quad f \in \mathcal{C} \tag{1.58}$$

with $\mathcal{C} \subset \mathcal{D}(A)$ a determining class.

1.5. Piecewise deterministic Markov processes

In Section 1.2 we constructed MPPs and it turns out that from any stochastic process mapping into $(G, \mathcal{G})$ it is possible to construct an MPP. This is done simply by letting

T_n be the time of the nth jump and $Y_n = X_{T_n}$ if the jump happens in finite time. A process is called piecewise deterministic (PDP) if one can reconstruct it from the induced MPP, that is its behavior between jumps is deterministic. The properties and construction of such processes—specifically finding a sufficient condition for a PDP to be Markovian—is the content of this section.

Regardless of the properties of a process it is clear that a collection of functions is needed if we at all should hope for $\mathbf{X}$ to be a PDP. Denoting the initial state of $\mathbf{X}$ by $x_0 \in G$, we assume that for any $n \in \mathbb{N}_0$ there exists a collection of G-valued functions $f^{(n)}_{z_n|x_0}(t)$ of z_n for $t \geq t_n$ if $t_n < \infty$ such that

$$X_t = f^{\langle t \rangle}_{Z_{\langle t \rangle}|x_0}(t), \qquad \text{where } \langle t \rangle := \overline{N}_t. \tag{1.59}$$

Hence by knowing the initial state, tracking the jump times and the intermediate marks up to t we know which of the functions describes X_t.

Before proceeding, we define what it means to be Markovian, where after Theorem 1.9 offers a sufficient condition for a process $\mathbf{X}$ to be Markovian.

The process $\mathbf{X}$ we wish to construct has to map into G and hence the mark space for the MPP $(\mathcal{T}, \mathcal{Y})$ is chosen such that $E = G$.[3] Naturally, the description involves the collection of $f^{(n)}_{z_n|x_0}$ as well as the kernels for the time points and marks. Not surprisingly, any sufficient condition constrains the transition probabilities for the time points as well as the mark space. Before proceeding it turns out that more can be said about the specific structure of the collection of the $f^{(n)}_{z_n|x_0}$'s.

We know that for $t \geq s$, $X_t = f^{\langle t \rangle}_{Z_{\langle t \rangle}|x_0}(t)$ must only depend on $X_s = f^{\langle s \rangle}_{Z_{\langle s \rangle}|x_0}(s)$, hence some function $\varphi_{st} : G \to G$ exists such that

$$\underbrace{f^{\langle t \rangle}_{Z_{\langle t \rangle}|x_0}(t)}_{=X_t} = \varphi_{st}\Big(\underbrace{f^{\langle s \rangle}_{Z_{\langle s \rangle}|x_0}(s)}_{=X_s} \Big). \tag{1.60}$$

Particularly, for t_k being the time point for the kth jump implying $f^{\langle t_k \rangle}_{Z_{\langle t_k \rangle}|x_0}(s) = y_k$ as a boundary condition, it holds that

$$f^{\langle t_k \rangle}_{Z_{\langle t_k \rangle}|x_0}(t) = \varphi_{t_k t}\left(y_k\right), \qquad t_k \leq t < t_{k+1}.$$

That is, the behavior of the process is determined by the a deterministic function up until the next jump. While this identity only holds between t_k and t_{k+1} it can be generalized to any $u \geq t$ as t_{k+1} is stochastic and it might even be that $t_{k+1} = \infty$. Using the identity, plugging into (1.60) and extrapolating to any time points $s \leq t \leq u$ with $y \in G$, we obtain

$$\varphi_{su}(y) = \varphi_{tu}(\varphi_{st}(y)),$$

implying that φ_{tt} is idempotent. From the current context it is seen that specifically $\varphi_{tt}(y) = y$ for all $y \in G$. Equivalently, summarizing

$$\varphi_{su} = \varphi_{tu} \circ \varphi_{st}, \qquad \varphi_{tt} = id, \tag{1.61}$$

where φ_{su}, $s \leq u$, can be thought of as the behavior from s to u. In the time-homogeneous case, (1.61) becomes particularly simple reducing to

$$\varphi_{s+t} = \varphi_s \circ \varphi_t, \qquad \varphi_0 = id, \tag{1.62}$$

[3] Assume that G, now serving as mark space, is a Borel space. In applications usually isomorphic to $\mathbb{R}^n$.

which shows that φ_s and φ_t commutes. Using φ_0 as neutral element the set $\{\varphi_t, t \geq 0\}$ with function composition as binary operator defines an Abelian semi group. Two examples of such functions are the step function $\varphi_s(x) = x$ and the exponential behavior with $\varphi_s(x) = xe^{Ks}$ for some $K \in \mathbb{R}$.

This concludes our analysis of the piecewise deterministic part of the process. Nearly ready to formulate Theorem 1.9 on the sufficient conditions for a PDP to be Markovian, we introduce some additional notation to aid handling the Markov kernels. The kernels use transition intensities $q_t(x, C)$ for $x \in \mathbb{R}_0$ and $C \in \mathcal{G}$. This intensity is decomposed into a part $q_t(x)$ triggering the jumps and a part $r_t(x, C)$ deciding where the process jumps to such that $q_t(x, C) = q_t(x)r_t(x, C)$. The conditions below are referred to as the *intensity conditions* .

- $q_t(x) := q_t(x, G)$ simply measures any jump from x at time t occurs as all marks are contained in G and it is referred to as the *total intensity*. We require that $(x, t) \mapsto q_t(x)$ is measurable and $\int_t^{t+h} q_s(\varphi_{ts}(x))ds < \infty$ for all (x, t) where $h(x, t) > 0$ can depend on (x, t). This is referred to as $q_t(x)$ being *right locally integrable*.

- We require that r_t for each t is a Markov kernel $(t, x) \mapsto r_t(x, C)$. It is seen as the conditional probability on G such that $r_t(x, C)$ is the probability that a jump leads to C given that it occurs from x at time t.

- We require $r_t(x, \{x\}) = 0$ where $t \in \mathbb{R}_0$ and $x \in G$ ensuring a genuine jump occurs.

At last, we are equipped to investigate the existence of piecewise deterministic Markov procecesses (PDMPs). The first part of the theorem describes the general case, whereas the second part focuses on the time-homogeneous case.

Theorem 1.9 (PDMP condition)**.**

- General case: *Let measurable $\varphi : \{s \leq t\} \times G \to G$ satisfy (1.61) with $t \mapsto \varphi_{st}(y)$ continuous on $[s, \infty[$ for all $s \in \mathbb{R}_0$ and $y \in G$. Further, suppose that q_t and r_t satisfy the intensity conditions. Then the PDP $\mathbf{X}$ given by*

$$X_0 = Y_0 \equiv x_0 \quad \text{and} \quad X_t = \varphi_{T_{\langle t \rangle}, t}(Y_{\langle t \rangle}) \tag{1.63}$$

is in fact a PDMP that does not depend on x_0. The Markov kernels are given by:

$$\overline{P}^{(0)}_{|x_0}(t) = \exp\left(-\int_0^t q_s(\varphi_{0s}(x_0))ds\right), \qquad t \in \mathbb{R}_0 \tag{1.64}$$

$$\pi^{(0)}_{t|x_0}(C) = r_t(\varphi_{0t}(x_0), C), \qquad t \in \mathbb{R}_0, C \in \mathcal{G}, \tag{1.65}$$

and

$$\overline{P}^{(n)}_{z_n|x_0}(t) = \exp\left(-\int_{t_n}^t q_s(\varphi_{t_n s}(y_n))ds\right), \tag{1.66}$$

$$\pi^{(n)}_{z_n, t|x_0}(C) = r_t(\varphi_{t_n t}(y_n), C), \qquad t \in \mathbb{R}_0, C \in \mathcal{G}, \tag{1.67}$$

with $n \in \mathbb{N}$, $t_1 < \ldots < t_n < t$, and $y_i \in G$.

- Time-homogeneous case: *A PDMP* $\mathbf{X}$ *satisfying the former requirements is* time-homogeneous *if in addition* q_t *and* r_t *do not depend on* t *and if there exists measurable* $\tilde{\varphi}$ *satisfying (1.62) with* $t \mapsto \tilde{\varphi}_t$ *continuous such that*

$$\tilde{\varphi}_{t-s} \equiv \varphi_{st}. \tag{1.68}$$

Note, that as φ and $\tilde{\varphi}$ operate on different spaces we usually write $\varphi : \mathbb{R}_0 \times G$ instead of $\tilde{\varphi}$ without running risks of ambiguity. Further, the Markov kernels from (1.66) and (1.67) in the time-homogeneous case reduce to:

$$\overline{P}^{(n)}_{z_n|x_0}(t) = \exp\left(-\int_0^{t-t_n} q(\varphi_s(y_n))ds\right), \tag{1.69}$$

$$\pi^{(n)}_{z_n,t|x_0}(C) = r(\varphi_{t-t_n}(y_n), C), \quad t \in \mathbb{R}_0, C \in \mathcal{G}. \tag{1.70}$$

A generalization of this theorem treating the case where the $P^{(n)}$ do not have densities exists and the reader is referred to Theorem 7.3.2 in [Ja06] p. 157.

Before proceeding, it is worth dwelling on a few of the observations established above. Despite the apparent complicated nature of the initial deterministic function $f^{\langle t\rangle}_{Z_{\langle t\rangle}|x_0}$, restricting attention to cases where the resulting process becomes Markovian simplifies matters immensely. Particularly in the time-homogeneous case we end up with a single function φ_s containing complete information about the behavior of the process between jumps. Further, Theorem 1.9 provides simple sufficient conditions to check. Finally, it should be noted that the expression for (1.66) and (1.67) as well as (1.69) and (1.70) do not dependent of x_0.

1.6. Renewal theory

The exposition of this section is primarily based on [KT75] with some stray ideas taken from [Ro83].

Before developing the rigorous mathematical theory a motivating example is offered describing stochastic processes who at stochastic times begin anew. As such they are on one hand a generalization of the Poisson process while on the other hand a special case of SPPs/MPPs.

Consider a light bulb placed at time $T_0 = 0$. The light bulb fails at some stochastic time T_1 whereupon it is immediately replaced with a second light bulb of the same make. This second light bulb fails after V_2, that is at time T_2, and is again replaced with an identical light bulb.

Formalizing the experiment, V_k is viewed as a strictly positive iid random variables with distribution F and, if it exists, density f such that $F(0) = 0$ and even $F(0^+) = 0$ measuring[4] the interarrival times between the kth and $k+1$th event. $T_n := \sum_{k=1}^n V_k$ is the time of the nth occurrence while $N_t := \sup\{n : T_n \leq t\}$ gives the number of event up to and including t is of central interest. The following relation holds

$$N_t \geq n \iff T_n \leq t. \tag{1.71}$$

The process $\mathbf{N} = (N_t)_{t\geq 0}$ is called a *renewal process*. The expectation of the interarrival times is also used extensively and we put $\mu := \mathrm{E}[V_1]$.

[4]The latter being equivalent to stating that there can be no explosions, that is we consider a genuine point process.

Remark It is apparent how the renewal theory fits into the point process framework described earlier. Our strategy in attacking jump-diffusions is to start a probabilistic replica of a given diffusion—corresponding to the deterministic pieces of the PDMP—at each renewal. This section shows that renewal processes can be made stationary and this is a cornerstone of the theory developed in the following chapters.

A key component of renewal theory is to study the properties of $N(t)$ depending on the distribution of the interarrival times F. Of particular interest is the *renewal function* given by

$$m(t) = \mathrm{E}[N_t]. \tag{1.72}$$

Knowing F it is easy to find a closed description of T_n's distribution, say F_n, namely the recurrence relation

$$F_n(x) = \int_0^x F_{n-1}(x-y)dF(y) = (F_{n-1} \star F)(x) = F^{n\star} \tag{1.73}$$

which is equivalent to $(F_{n-1} * f)(x)$ if F has a density (see (A.3) and (A.4) in Section A.2 for the definition of "$*$" and "$\star$."). Using (1.71) it follows that $\mathbb{P}(N_t \geq k) = F_k(t)$ and therefore

$$m(t) = \sum_{k=1}^{\infty} \mathbb{P}(N_t \geq k) = \sum_{k=1}^{\infty} F_k(t), \tag{1.74}$$

where the first equality follows from (1.2). We wish to show that the sum converges and is finite. From (1.73) it is seen that

$$F_n = (F_{n-k} \star F_k)(t) \leq F_{n-k}(t)F_k(t)$$

as all F_k's are distribution function and therefore increasing. In particular for fixed r it is seen that $F_{nr+k}(t) \leq (F_r(t))^n \cdot F_k(t)$ for $0 \leq k \leq r-1$ such that

$$m(t) = \sum_{k=1}^{\infty} F_k = \left(\sum_{k=1}^{r-1} F_k\right) \cdot \sum_{n=0}^{\infty} (F_r(t))^n, \quad t > 0$$

where the first sum is finite and the second sum is a geometric series converging if and only if $F_r(t) < 1$. As we have assumed $F(0^+) = 0$ (no explosions) it follows that for every t some r is large enough such that this is true and we conclude that the sum, and hence the renewal function, is finite.

Other associated random variables of interest are the *current life* δ_t and *excess lifetime* γ_t at time t. These are given by

$$\delta_t := t - T_{N_t} \quad \text{and} \quad \gamma_t := T_{N_t+1} - t. \tag{1.75}$$

The intuition is that δ_t looks "backward" and measures the current life of the light bulb, while γ_t looks "forward" and gives the remaining lifetime. As a side note, their sum is quite naturally called the *total life* and denoted by β_t.

1.6.1. Renewal equations and the elementary renewal theorem

Having the initial setup in place we are now ready to prove some interesting results and first wish to prove

$$m(t) = F(t) + \int_0^t m(t-y)dF(y) = F(t) + (m \star F)(t), \quad t \geq 0. \tag{1.76}$$

Begin by considering

$$\mathrm{E}[N_t|T_1 = s] = \begin{cases} 0 & \text{if } s > t \\ 1 + m(t-s) & \text{if } s \leq t \end{cases}$$

which is true since the first event happens at time s then N_t must necessarily be 0 for all $t < s$ and so must the expectation of N_t. If instead the first event has already occurred then the probability for another event occurring must be $m(t-s)$ as the time between events are iid. Arguments of this type are ubiquitous in renewal theory and referred to as *renewal arguments*. Using the law of total probability, we proceed to find

$$m(t) = \int_0^\infty \mathrm{E}[N_t|T_1 = s]dF(s) = \int_0^t 1 + m(t-s)dF(s) = F(t) + (m \star F)(t)$$

as desired. Equations of the type (1.76) are called *renewal equations* and are treated in the following theorem.

Theorem 1.10. *If g is a bounded function there exists a unique function G, bounded on finite intervals, and satisfying*

$$G(t) = g(t) + \int_0^t G(t-y)dF(x) = g(t) + (G \star F)(t). \tag{1.77}$$

This G is

$$G(t) = g(t) + \int_0^t g(t-x)dm(x) = g(t) + (g \star m)(t) \tag{1.78}$$

with m the renewal function corresponding to F.

Proof. Existence and uniqueness needs to be proved. Considering the existence of the solution first it is shown that G in (1.78) actually solves (1.77) and is bounded on finite intervals.

Using (1.74) it follows that

$$\begin{aligned} G(t) &= g(t) + (g \star m)(t) = g(t) + (g \star F)(t) + (g \star \sum_{k=2}^{\infty} F_k)(t) \\ &= g(t) + \Big((g + g \star \sum_{k=1}^{\infty} F_k) \star F\Big)(t) = g(t) + (G \star F)(t), \end{aligned}$$

that is G solves (1.77). To check that G is indeed bounded on finite intervals $[0, T]$ observe

$$\sup_{0 \leq t \leq T} |G(t)| \leq \sup_{0 \leq t \leq T} |g(t)| \Big(1 + \int_0^T dm(t)\Big) < \infty$$

since g is assumed bounded and the renewal function is finite as well.

For uniqueness of the solution observe that as G solves (1.77) it inductively follows that

$$G = g + (g \star F + (G \star F) \star F) = \ldots = g + (g \star \sum_{k=1}^{n-1} F_k) + (G \star F_n)$$

By assumption g is bounded and as $m(t) = \sum F_k$ it follows that $g \star m(t)$ is an integrable upper bound for $(g \star \sum_{k=1}^{n-1} F_k)$ the latter therefore converging to $(g \star m)$

for $n \to \infty$. It thus suffices to show that $\lim_{n\to\infty} |(G \star F_n)(t)| = 0$ for fixed t. Hence consider

$$|(G \star F_n)(t)| \leq F_n(t) \cdot \sup_{0\leq x\leq t} G(x)$$

Since it has been shown that $\lim_{n\to\infty} F_n(t) = 0$ and as G is bounded on finite intervals implying that the above tends to 0 for $n \to \infty$. It has therefore been shown that any G bounded on finite intervals satisfying (1.77) must be of the form (1.78). □

On first sight it might seem a rather peculiar idea to study equations like (1.77). Let us for a moment imagine that we can find a G of interest such that (1.77) is fulfilled with $g \equiv K$, a constant function. Plugging into (1.78) would give the relation $G(t) = K(1 + m(t))$ thus linking G to the renewal function in a very simple way. Depending on G this approach is potentially useful for exploring properties of the renewal function.

As a specific application of this idea let $G(t) = \mathrm{E}[T_{N_t+1}]$. A renewal argument conditioning on the time of the first jump gives

$$\mathrm{E}[S_{N_t+1}|T_1 = s] = \begin{cases} s & \text{if } t < s \\ s + G(t-s) & \text{if } t \geq s \end{cases}$$

such that the law of total probability gives

$$\begin{aligned} G(t) &= \int_0^\infty \mathrm{E}[S_{N_t+1}|T_1 = s]dF(s) = \int_0^\infty sdF(s) + \int_0^s G(t-s)dF(s) \\ &= \mu + (G \star F)(t) \end{aligned}$$

remembering $\mu = \mathrm{E}[V_1]$ and we conclude

$$\mathrm{E}[T_{N_t+1}] = \mu(1 + m(t)). \tag{1.79}$$

This is used to prove the following theorem.

Theorem 1.11 (Elementary Renewal Theorem)**.** *Let $m(t)$ be the renewal function linked to $V_1, \ldots$ positive iid random variables with $\mu = \mathrm{E}[V_1] < \infty$. Then*

$$\lim_{t\to\infty} \frac{m(t)}{t} = \frac{1}{\mu}. \tag{1.80}$$

Intuition Returning to the light bulb example the interpretation of (1.80) is that in the long run we would expect the average rate at which to change the light bulbs to equal the average life time of a light bulb. Thus (1.80) can be thought of as a sort of law of large numbers for renewal processes.

Proof. First it is proved that

$$\lim_{t\to\infty} \inf \frac{m(t)}{t} \geq \frac{1}{\mu}.$$

This easily follows from the fact that $t < S_{N_t+1}$ implying by (1.79) that $t < \mu(1 + m(t))$ such that

$$\frac{m(t)}{t} > \frac{1}{\mu} - \frac{1}{t}.$$

To prove the opposite inequality is more involved and an auxiliary truncated renewal process is needed. Define for arbitrary $C > 0$

$$V_i^C = \begin{cases} V_i & \text{if } V_i < C \\ C & \text{if } V_i \geq C \end{cases}.$$

Note, that V_i^C is in fact a genuine renewal process as the time between renewals are strictly positive and iid. Hence T_n^C, N_t^C, $m^C(t)$ and μ^C are all well-defined and (1.79) gives

$$\mathrm{E}[S^C_{N_t^C+1}] = \mu^C(1 + m^C(t)). \tag{1.81}$$

As the time between events for the truncated process can be at most C it follows that $t + C \geq S^C_{N_t^C+1}$. As furthermore $T_i^C \leq T_i$ it must be that $m^C(t) \geq m(t)$ and thus from (1.81) it is concluded that

$$t + C \geq \mu^C(1 + m(t)).$$

Rearranging and taking the limsup gives

$$\limsup_{t\to\infty} \frac{m(t)}{t} \leq \limsup_{t\to\infty} \frac{1}{\mu^C} + \frac{1}{t}\left(\frac{C}{\mu^C} - 1\right) = \frac{1}{\mu^C} \tag{1.82}$$

Unsurprisingly by monotone convergence using (1.1),

$$\mu^C = \int_0^\infty 1_{(0,C)}(t)x + 1_{(C,\infty)}(t)C dF(x) \to \int_0^\infty x dF(x) = \mu$$

for $C \to \infty$. By (1.82)

$$\limsup_{t\to\infty} \frac{m(t)}{t} = \lim_{C\to\infty} \limsup_{t\to\infty} \frac{m(t)}{t} \leq \lim_{C\to\infty} \frac{1}{\mu^C} = \frac{1}{\mu}$$

establishing the opposite inequality and concluding the proof. □

1.6.2. Key renewal theorem

We are now ready to formulate the key renewal theorem crucial for the models we build in Chapter 2 and 3. Before stating the theorem itself we need the following definitions

Definition 1.10. A distribution is said to be *arithmetic* if the only sets of positive measure are in $\{0, \pm j, \pm 2j, \ldots\}$. The largest such j that captures all values of the distribution is called the *span* of the distribution.

Thus a continuous distribution is not arithmetic while a discrete distribution with possible values in $0, 1, \ldots$ is arithmetic with span 1. We only use non-arithmetic distributions in this thesis.

Definition 1.11. A function $g : [0, \infty) \to \mathbb{R}$ is said to be *directly Riemann integrable* if for any $\delta > 0$ and $n = 1, 2, \ldots$

$$\underline{s} = \delta \sum_{n=1}^{\infty} \underline{m}_n \quad \text{and} \quad \overline{s} = \delta \sum_{n=1}^{\infty} \overline{m}_n \tag{1.83}$$

both converge absolutely and $\overline{s} - \underline{s} = 0$ for $\delta \to 0$ where

$$\underline{m}_n := \min\{g(t), (n-1)\delta \leq t \leq n\delta\},$$
$$\overline{m}_n := \max\{g(t), (n-1)\delta \leq t \leq n\delta\}.$$

Note that this is equivalent to g being Riemann integrable on every compact set and either

$$\sum_{n=0}^{\infty} \sup\{|g(t)| : n \leq t \leq n+1\} < \infty \quad \text{or} \quad \int |g| dt < \infty.$$

Theorem 1.12 (Key Renewal Theorem)**.** *Let F be the distribution function of a positive random variable with finite mean μ. Suppose g is directly Riemann integrable and G solves*

$$G(t) = g(t) + (G \star F)(t).$$

- *i) If F is not arithmetic then*

$$\lim_{t\to\infty} G(t) = \begin{cases} \frac{1}{\mu}\int_0^\infty g(x)dx & \text{if } \mu < \infty, \\ 0 & \text{if } \mu = \infty. \end{cases} \tag{1.84}$$

- *ii) If F is arithmetic with span j, then for all $C > 0$,*

$$\lim_{n\to\infty} G(C + nj) = \begin{cases} \frac{j}{\mu}\sum_{l=0}^\infty g(C + lj) & \text{if } \mu < \infty, \\ 0 & \text{if } \mu = \infty. \end{cases} \tag{1.85}$$

Proof. See [Fe66] p. 346-351. Theorem 2 states the result in the above notation and is proved to be equivalent to Theorem 1. □

Although not apparent at first the theorem generalizes the elementary renewal theorem. An equivalent formulation is presented in [KT75] p. 192 more suitable for that purpose, but instead a few quick calculations and application of the above formulation gives us the desired generalization.

We wish to show that for arbitrary $h > 0$

$$\lim_{t\to\infty} m(t+h) - m(t) = \frac{h}{\mu} \tag{1.86}$$

which intuitively can be thought of as that the number of renewals over an interval of length h is h/μ as long as the renewal process has been going for a long time, that is t tends to infinity.

Consider $g(t) = 1_{(0,h)}(t)$. Theorem 1.10 implies that $G(t) = m(t+h) - m(t)$ for $t > h$ and if F is non-arithmetic the key renewal theorem gives

$$\lim_{t\to\infty} m(t+h) - m(t) = \lim_{t\to\infty} G(t) = \frac{h}{\mu}$$

with the convention that $h/\mu = 0$ if $\mu = \infty$. If F instead is arithmetic an analog argument shows that (1.86) continues to hold as long as h is a multiple of the span j for F.

This does not really imply the elementary renewal theorem yet, but is rather an "infinitesimal" or "differential" version if one will. It can though be used to approximate and establish $m(t) \sim t/\mu$.

Distribution of excess lifetime There are many applications of the key renewal theorem and particularly relevant for what follows is that it allows us to find the limiting distribution of the excess lifetime γ_t defined in (1.75) as $t \to \infty$. Let $G_z(t) = \mathbb{P}(\gamma_t > z)$ and use a renewal argument to find

$$\mathbb{P}(\gamma_t > z | T_1 = s) = \begin{cases} 1 & \text{if } s > t + z, \\ 0 & \text{if } t + z \geq s > t, \\ G_z(t - s) & \text{if } t \geq s > 0. \end{cases}$$

By the law of total probability

$$\begin{aligned} G_z(t) &= \int_0^\infty \mathbb{P}(\gamma_t > z | T_1 = s)\, dF(s) = \int_{t+z}^\infty dF(s) + (G_z \star F)(t) \\ &= 1 - F(t + z) + (G_z \star F)(t). \end{aligned}$$

We wish to apply the key renewal theorem in order to find the limiting distribution for $t \to \infty$ and hence we must assume that $1 - F(t + z)$ is integrable (on the positive axis). Put $g(t) = 1 - F(t + z)$ then g is continuous, bounded, and Riemann integrable on every compact set. Assuming $\mu < \infty$

$$\int_0^\infty |g(t)|\, dt = \int_z^\infty 1 - F(x)\, dx \leq \mu$$

shows that g is DRI. The key renewal theorem now yields for F non-arithmetic

$$\lim_{t\to\infty} \mathbb{P}(\gamma_t > z) = \mu^{-1} \int_z^\infty 1 - F(x)\, dx, \qquad z > 0. \tag{1.87}$$

Similarly using renewal arguments, the limiting distribution of the current life δ_t and total life β_t can be found to be

$$\lim_{t\to\infty} \mathbb{P}(\delta_t > z) = \mu^{-1} \int_z^\infty 1 - F(y)\, dy = \lim_{t\to\infty} \mathbb{P}(\gamma_t > z) \tag{1.88}$$

and with F_β the limit distribution function of the total life

$$F_\beta(z) := \lim_{t\to\infty} \mathbb{P}(\beta_t \leq z) = \mu^{-1} \int_0^z y\, dF(y). \tag{1.89}$$

In particular it is important to note that the limiting distribution for the excess lifetime and current life are identical.

This might be surprising and we offer two heuristic arguments.

First, consider a double infinite renewal process, that is a renewal process over the entire real axis such that T_n is defined for all $n \in \mathbb{Z}$. If the limiting distribution is chosen such that this process is stationary the distribution of the excess lifetime is the same for all t. But the reversed of this renewal process (that is "played backwards") is again a stationary renewal process with the same excess lifetime distribution. By reversing the process we switch the meaning of excess lifetime and current life; thus the distribution of these must be equal.

Another argument is to fix a certain time point $0 \ll t_0$ such that the process has been running for a very long time and is close to its limiting distribution. We inspect the renewal process at t_0 and observe that

probability of renewal at $t_0 - s$ equals probability of renewal at $t_0 + s$

as the interarrival times are iid. The left-hand side event is the current life time being s while the right-hand side event is the excess life time being s. This gives that the distribution of the current and excess life time must both equal the limit distribution.

An interesting peculiarity is the *inspection paradox*.[5] Consider the mean of the total life, which by (1.89) is given as

$$\mathrm{E}[\lim_{t\to\infty} \beta_t] = \int_0^\infty z\, dF_\beta(z) = \mu^{-1} \int_0^\infty z^2\, dF(z) \geq \mu^{-1} \left(\int_0^\infty z\, dF(z) \right)^2 = \mu$$

since $\mathrm{E}[V_1^2] \geq \mathrm{E}[V_1]^2$ (non-negativity of variance or Cauchy-Schwarz inequality) with strict inequality prevailing unless F is degenerate. That is the mean of the total life is strictly larger than the mean interarrival time. Or, in everyday terms, if one decides to measure the current age and the remaining lifetime at a specific time point t it is on average strictly larger than the average time between renewals.

The elementary renewal theorem being a sort of law of large numbers we mention that an analogy to the central limit theorem exists as well. The proof is short, see [KT75] p. 205.

1.6.3. Stationary renewal processes

So far we have assumed that the replacement experiment starts immediately with all the V_i's iid, but it is reasonable to allow V_1 to follow a different distribution. Think of the light bulb example where the first light bulb has already been in place for some time and therefore, even though of the same make as the following, it has a different distribution. Such a process is called a *delayed renewal process* and we denote the distribution function of V_1 by H still independent of $V_2, \ldots$ In contrast a 0-delayed renewal process may be prefixed with *pure* and it is seen that a renewal process is pure if and only $T_0 = 0$.

We associate a renewal function $m^D(t)$ with the delayed renewal process and in general $m^D(t) \neq m(t) = \sum F_k(t)$. As we only mangle with the first distribution one would think that the asymptotic properties of the renewal function stays the same, which is indeed the fact and for example

$$\lim_{t\to\infty} m^D(t+h) - m^D(t) = \frac{h}{\mu}, \quad h > 0$$

holds.

Now what if we choose $H(x) := \mu^{-1} \int_0^x 1 - F(y) dy = \mathbb{P}(\gamma_t \leq x)$? This means we are starting the renewal process in such a way that the first light bulb fails according to the limiting distribution of the excess lifetime and one would expect some nice stationary properties. We show that for this delayed renewal process

$$m^D(t) = \frac{t}{\mu} \quad \text{and} \quad \mathbb{P}(\gamma_t^D \leq x) = H(x), \tag{1.90}$$

that is, the results we know to hold asymptotically are indeed identities and in particular $\mathbb{P}(\gamma_t^D \leq x)$ is independent of t. It therefore makes sense to call such a renewal process stationary with stationary distribution H. It can be shown that $F(x) = H(x)$ characterizes the Poisson process among renewal processes.

[5] Sometimes "waiting time paradox."

Using a well-known renewal argument $\mathrm{E}[N_t^D|T_1 = s] = 1_{s<t}(t)(1 + m(t-s))$ used to show $m^D(t) = H(t) + (H \star m)(t)$, that is the unique solution of the renewal equation

$$m^D(t) = H(t) + (m^D \star F)(t).$$

The solution is unique and it therefore suffices to check that $m^D(t) = t/\mu$ satisfies the equation above.

$$H(t) + (m^D \star F)(t) = \mu^{-1}t + \mu^{-1}\left(-\int_0^t F(x)\,dx + \int_0^t t - x\,dF(x)\right) = \mu^{-1}t$$

since the factor in the braces equals zero which is seen using integration by parts

$$-\int_0^t F(x)\,dx = -tF(t) + \int_0^t x\,dF(x) = -t\int_0^t dF(x) + \int_0^t x\,dF(x).$$

Showing $\mathbb{P}(\gamma_t^D \leq x) = H(x)$ is a bit more lengthy and therefore omitted here, see [KT75] p. 200.

We wish to build a bit of intuition around this and a generic limiting argument might shed some light on the above result. If the renewal process has a limiting distribution and we start the process with this particular distribution it heuristically amounts to considering the original process for t_0 very large and start observing from there. In other words, by choosing the limiting distribution to begin with we have ensured that the convergence has already taken place and as the convergence also holds for the delayed process "we must already have arrived" at the limiting distribution.

1.6.4. Regenerative processes

Finally, we briefly mention the vast class of *regenerative processes* first mentioned by Feller [Fe66] p. 365 (discrete time only) and later made more precise by and investigated in depth by [Mi72]. They are of tremendous importance to this thesis and ubiquitous in Chapters 2 and 3.

Generally, a regenerative process is a process where with probability one there exists a time T_1 such that the process is started anew (regenerated) with a probabilistic replica of itself. This implies that time points $T_2, T_3, \ldots$ exist. The 'pieces' between the time points might be of virtually any nature and in case of diffusions a regenerative process is a jump-diffusion.

A regenerative process is stationary if the resetting times are governed by a stationary renewal process. We only give a heuristic argument for this fact considered well-known in the literature. Consider two points s and t and the distribution of the process at these times H_s and H_t. As the pieces of the regenerative process are probabilistic replicas H_s and H_t can only be different if the distribution of the current age of the regenerative piece containing s and t is different. But with the underlying renewal process being stationary the current lifetime has the same distribution at all times. Hence H_s and H_t must describe the same distribution.

CHAPTER 2

Jump-diffusion models using jump targets

This chapter deals with finding the stationary distribution of a particular class of jump-diffusion models. It is characterized by iid jumps independent of the position of the process, while the process between jumps is a diffusion. The particulars of the setup is described in Section 2.1. That a unique stationary distribution exists for a large subset is shown in Section 2.2.

In some cases one can find an explicit expression for the stationary initial distribution. Considering the Brownian motion as the diffusion part, the stationary distribution is found in Section 2.3 using renewal theory, while Section 2.4 and 2.5 uses the infinitesimal generator. Section 2.6 formulates a criterion for diffusions in general that any stationary distribution must satisfy and Section 2.7 treats diffusions which themselves have stationary initial distributions.

2.1. The model

Loosely speaking the process $\mathbf{X} = (X_t)_{t\geq 0}$ is pieced together using parts of diffusions $\mathbf{X}^{(i)}$ between jump times T_{i-1} and T_i for $i \in \{1, 2, \ldots\}$. The following model is referred to as a *jump-diffusion with jump targets*.

The ith jump target has distribution Y_i, $i \in \{0, 1, \ldots\}$, where Y_0 is the boundary condition for the very first diffusion. The Y_i's are iid for all i. They are distributed according to G and with density, when it exists, given by g.

Let each $X^{(i)}$ be a diffusion with fixed measurable functions b and σ such that each $X^{(i)}$ solves the SDE

$$dX_t^{(i)} = b(X_t^{(i)})dt + \sigma(X_t^{(i)})dB_t, \qquad X_0^{(i)} = Y_{i-1}. \tag{2.1}$$

Thus the $\mathbf{X}^{(i)}$'s are probabilistic replica of the same diffusion if and only if the Y_i's follow the same, possible degenerate, distribution for all i. These make up the continuous parts of the jump-diffusion $\mathbf{X}$.

The jumps are comprised of two ingredients: i) the target distribution of the jump, and ii) a point process triggering the jumps.

The jump times are denoted by $T_0, T_1, T_2, \ldots$ with $T_0 = 0$. They are chosen such that the interarrival times $T_1 - T_0, T_2 - T_1, \ldots$ are independent and from the second onwards distributed according to an intensity $\lambda(x) > 0$. The resulting distribution

function for the interarrival times is denoted by F and on occasion by F_G when the distribution of Y_i needs to be specified.

The interarrival times $V_i := T_i - T_{i-1}$, $i \geq 1$, from the second jump onwards are identically distributed with cumulative distribution function F as the diffusion pieces are probabilistic copies. We assume that $\mathbb{P}_G(T_1 < \infty) = 1$ such that a jump in finite time guaranteed. Equivalently the assumption is often stated as $\mu := \mathrm{E}_G[V_i] = \mathrm{E}_G[T_1] < \infty$. Often jumps from a Poisson process are considered in which case G becomes superfluous in the notation $\mathrm{E}_G[T_1]$.

It is clear that when conditioning on $X^{(1)}$ with $X_0^{(1)} = Y_0$ we obtain

$$\mathbb{P}_G(T_1 > t | X^{(1)}) = \exp\left(-\int_0^t \lambda(X_s^{(1)})ds\right). \tag{2.2}$$

After the first jump a new and independent (of everything) copy of the diffusion $X^{(2)}$ is created with starting point determined by Y_1, that is $X_0^{(2)} = Y_1$. Obviously, it therefore suffices to condition on $(X_t^{(1)})_{0 \leq t \leq T_1}$ in (2.2). Having knowledge of $X^{(2)}$ and the initial value Y_1 gives information about the second waiting time, namely

$$\mathbb{P}(V_2 > t | X^{(2)}, Y_1) = \exp\left(-\int_0^t \lambda(X_s^{(2)})ds\right).$$

Continuing inductively, we can glue together a jump-diffusion $\mathbf{X}$ described by

$$X_t = X_{t-T_i}^{(i)} \quad \text{for } T_i \leq t < T_{i+1}, \quad i \in \{0, 1, \ldots\}. \tag{2.3}$$

$\mathbf{X}$ is Markovian with generator A given by

$$\begin{aligned} Af(x) &= b(x)f'(x) + \frac{\sigma^2(x)}{2}f''(x) + \lambda(x)\int f(y) - f(x)dG(y) \\ &= A^d(f) + \lambda(x)\int f(y) - f(x)dG(y) \end{aligned} \tag{2.4}$$

for $f \in C^2(\mathbb{R})$ where A^d denotes the usual generator for the diffusion.

Remark More generally the generator for a jump-diffusion A is for $f \in C^2(\mathbb{R})$ given as

$$Af = A^d f + \lambda(x)\int f(y) - f(x)dG_x(y). \tag{2.5}$$

Note that y is the target of the jump while Y_i's distribution is allowed to depend on x, the position of the jump diffusion at the jump time.

2.2. Stationary distributions

When we are searching for stationary distributions it is important to note that the jump-diffusions consist of three parts:

1. the distribution of the diffusions $X^{(i)}$,
2. the distribution of the jump targets Y_i, and
3. the distribution of the underlying point process T_i.

These have to be chosen in a way such that the resulting process turns stationary.

Certain mild conditions are required for a stationary distribution to exists. On average, the first jump must take place in finite time and the average time between jumps must be finite. The following theorem makes the statement precise.

Theorem 2.1. *Consider a jump-diffusion with jump targets model where the first diffusion is started according to Y_1, that is $X_0^{(1)} \stackrel{\mathcal{D}}{=} Y_1$. If $\mu = \mathrm{E}_G[T_1] < \infty$ then a stationary distribution F_0 exists. F_0 satisfies the following relation*

$$\int f(x)dF_0(x) = \frac{1}{\mu}\,\mathrm{E}_G\left[\int_0^{T_1} f(X_s)ds\right], \qquad f \in C_b(\mathbb{R}) \tag{2.6}$$

and the right-hand side of (2.6) defines a probability measure.

Proof. Recall that G is the distribution of Y_i. The idea of the proof is to consider

$$R(t) := \mathrm{E}_G[f(X_t)] \text{ and show it converges to } \int f(x)dF_0(x) \text{ as } t \to \infty$$

for a class of f's sufficiently large to determine the distribution of X_t. This is done using the Key Renewal Theorem (Theorem 1.12) where after uniqueness and existence are shown.

Splitting after whether the first jump occurs before or after t and using a renewal argument (recall that F_G is the distribution of the waiting times when the jump targets have distribution G), one gets

$$\begin{aligned}
R(t) &= \mathrm{E}_G[f(X_t); T_1 \le t] + \mathrm{E}_G[f(X_t); T_1 > t] \\
&= \int_0^t \mathrm{E}_G[f(X_t); T_1 = s]\, dF_G(s) + \underbrace{\mathrm{E}_G[f(X_t^{(1)}); T_1 > t]}_{=:g(t)} \\
&= g(t) + \int_0^t R(t-s)\, dF_G(s) \\
&= g(t) + (F * R)(t), \quad t > 0
\end{aligned}$$

to which the Key Renewal Theorem is applied. The distribution of T_1, F, is non-arithmetic as $\lambda(x) > 0$.

One needs to determine a set of f's such that g is directly Riemann-integrable (DRI) and the following shows that $f \in C_b(\mathbb{R})$ is such a set. Being Riemann integrable on every compact set with $\int |g|dt < \infty$ is equivalent to being DRI. Rewrite g as

$$\begin{aligned}
g(t) &= \mathrm{E}_G\left[\mathrm{E}_G[f(X_t^{(1)}); T_1 > t]\Big| X^{(1)}\right] = \mathrm{E}_G[f(X_t^{(1)})] \cdot \mathbb{P}(T_1 > t | X^{(1)}) \\
&= \mathrm{E}_G[f(X_t^{(1)})] \cdot \exp\left(-\int_0^t \lambda(X_s^{(1)})\, ds\right)
\end{aligned}$$

from which it follows that if $f(x)$ is continuous then so is $g(t)$. Let furthermore f be bounded by say K then

$$|g(t)| \le \mathrm{E}_G[|f(X_t)|; T_1 > t] \le K\,\mathrm{E}_G[1; T_1 > t] \le K,$$

that is, g is bounded by K as well. Thus g is a continuous and bounded function and such functions are Riemann integrable on every compact set. That $\int |g|\, dt < \infty$ follows from f being bounded since

$$\int_0^\infty |g|\, dt \le \int_0^\infty \mathrm{E}_G[|f(X_t)|; T_1 > t]\, dt = \mathrm{E}_G\left[\int_0^{T_1} |f(X_t)|\, dt\right] \le K\,\mathrm{E}_G[T_1] < \infty.$$

Hence the Key Renewal Theorem can be applied to such f's. In addition, for these f's the interchange of the order of integration in the following calculation is justified.

As $\mu < \infty$ the Key Renewal Theorem is applied to find the limit

$$\begin{aligned}\lim_{t\to\infty} R(t) &= \frac{1}{\mu}\int_0^\infty g(s)ds = \frac{1}{\mu}\int_0^\infty \mathrm{E}_G[f(X_s); T_1 > s]ds \\ &= \frac{1}{\mu}\int_0^\infty \int_{T_1^{-1}(s,\infty)} f(X_s(\omega))\, P(d\omega)ds = \frac{1}{\mu}\int_\Omega \int_0^{T_1} f(X_s(\omega))\, dsP(d\omega) \\ &= \frac{1}{\mu}\,\mathrm{E}_G\left[\int_0^{T_1} f(X_s)ds\right]\end{aligned}$$

Denote the right-hand side by $\Lambda(f)$ and note that it does not depend on the starting position x_0 of the jump-diffusion. While the limit only holds for certain f, $\Lambda(f)$ is well-defined for any measurable f. $\Lambda(1_{\mathbb{R}}) = 1$ and for $A_1, A_2, \ldots$ pairwise disjoint $\Lambda(\sum 1_{A_i}) = \sum \Lambda(1_{A_i})$ by dominated convergence and thus Λ defines a genuine probability.

Thus it has been shown that Λ defines a probability and that $R(t)$ converges to this for sufficiently many f in order to determine $\lim_{t\to\infty} R(t)$ uniquely.

To argue that a stationary distribution is determined uniquely by this has two parts. First it must be checked if it actually defines a probability measure and secondly that it is actually stationary.

Let ν_G^t denote the probability measure of X_t at time t when $X_0 \overset{\mathcal{D}}{=} Y_i$ then

$$\int f(x)d\nu_G^t(x) = \mathrm{E}_G[f(X_t)] = R(t) \to \int f(x)dF_0(x) \qquad \text{for } t \to \infty$$

for $f \in C_b(\mathbb{R})$. This per definition implies $\nu_G^t \xrightarrow{w} F_0$. What remains is to show that F_0 is a stationary distribution and that it is the only one.

That it is the stationary distribution follows from

$$F_0(f) = \lim_{t\to\infty} \nu_G^{t+s}(f) = \lim_{t\to\infty} \int \nu_y^s(f)d\nu_G^t(y), \qquad s > 0$$

which holds because $v_y^s(f) \in C_b(\mathbb{R})$ for the diffusions considered here. Hence

$$F_0(f) = \int \nu_y^s(f)dF_0(y) = \nu_{F_0}^s(f)$$

for $s > 0$ and $f \in C_b(\mathbb{R})$.

The uniqueness of the stationary distribution relies on the fact that the underlying diffusion is regular, that is attains all values throughout the interval it lives on. No formal proof is offered though. □

Theorem 2.1 is an existence/uniqueness theorem in the most classical sense: It offers neither a construction nor much explicitness, but it is soothing to know that the objects we are looking for actually exists. Further, if we are able to derive the stationary distribution in more than one way we know they must be the same.

Further note that one assumed $\mu < \infty$ such that jumps are guaranteed to occur in "reasonable" time. While this is trivially true for a Poisson process it is not necessarily so for arbitrary $\lambda(x)$. We could imagine a diffusion twist around a level for which the intensity is minute implying that no jumps occur or at very least that $\mu = \infty$. A non-constant intensity, as we shall see, complicates matters considerably. Nevertheless, if it is bounded below Theorem 2.1 immediately gives:

Corollary 2.2. *Consider a jump-diffusion with jump targets attaining values in* $]l, r[$ *and intensity* $\lambda(x)$*. If*

$$\lambda(x) > \lambda_0 > 0 \qquad \text{for all } x \in]l, r[$$

then a stationary distribution exists.

Proof. If one can show that $\mu = \mathrm{E}_G[T_1] < \infty$ then Theorem 2.1 gives the desired. By assumption it follows that

$$\begin{aligned}\mathrm{E}_G[T_1] &= \int_0^\infty \mathbb{P}_G(T_1 > t|X^{(1)})dt = \int_0^\infty \exp\left(-\int_0^t \lambda(X_s^{(1)})ds\right)dt \\ &< \int_0^\infty \exp(-\lambda_0 t)dt = \frac{1}{\lambda_0}.\end{aligned}$$

□

The intuition of the above corollary is that the intensity being bounded below by a constant λ_0 implies that regardless the behavior of $\mathbf{X}$ the jumps are always more likely to occur than for a Poisson process with intensity λ_0.

What follows is a collection of specific setups for which the stationary distribution can be determined explicitly.

2.3. Description of stationary density via renewal theory

In the preceding section we proved that a unique stationary distribution exists under very general conditions. While the proof was based on the Key Renewal Theorem, renewal theory can also be used to find an explicit expression for the density.

Consider a stationary renewal process which exists provided $\mu < \infty$. Recall that, at time t the distribution of the current life time δ_t equals that of the current life time at any other time, say s. The same holds for the excess life time γ_t and it even holds that $\delta_t \overset{\mathcal{D}}{=} \gamma_t$, $t > 0$.

For the jump-diffusion with a stationary point process this implies that at time t the current life of the diffusion has the same distribution as at any other time point.

Two probabilistic replica of the diffusion with boundary condition given by Y_i have the same distribution for any specific t. At time t, the jump-diffusion's diffusion part has been alive for δ_t, which has a stationary distribution with density f_ρ. Thus the stationary distribution for the jump-diffusion must be that of $X_{\delta_t}^{(i)}$. If densities exist for the transition probabilities the stationary density u for the jump-diffusion is given by

$$u(x) = \int dG(y) \int dt\, p_t(x, y) f_\rho(t). \tag{2.7}$$

Note that this line of reasoning does not depend on $\mathbf{X}^{(i)}$'s being diffusions, but merely that they have a transition probability. On the other hand this approach shows that the stationary distribution for jump-diffusion requires the underlying point process to be stationary and as, by Theorem 2.1, the stationary distribution is unique we formulate the following corollary.

Corollary 2.3. *For any jump-diffusion with independent jump targets governed by a stationary distribution the underlying point process must follow its stationary distribution.*

2.4. Brownian motion returning to 0

Consider the special case of $g \equiv \varepsilon_0$, that is at every jump we return to 0. Further, the diffusion considered is the Brownian motion and the jump-intensity is constant $\lambda(x) \equiv \lambda > 0$ such that the jump process becomes a Poisson process. We begin the analysis by considering a standard Brownian motion and then generalize to an arbitrary Brownian motion.

We show two ways of getting the stationary distribution in this setup. The first approach is inspired by renewal theory and uses (2.7). In the case of a Poisson process this becomes particularly simple with $f_\rho(t) = \lambda \exp(-\lambda t)$, that is the Poisson distribution itself. The second approach uses the infinitesimal generator to derive a differential equation which any stationary distribution must solve.

The results of this section are summarized in Theorem 2.4.

2.4.1. Standard Brownian motion

Recall from (1.37) that the transition probability for a standard Brownian motion is given by

$$p_t(0,x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right).$$

We quest the stationary distribution, its density denoted by u, and relation (2.7) yields

$$u(x) = \int_0^\infty p_t(0,x) f_\rho(t)\, dt = \frac{\lambda}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{t}} \exp\left(-\frac{x^2}{2}\frac{1}{t} - \lambda t\right) dt, \quad x \in \mathbb{R}. \tag{2.8}$$

We shall need the first derivative of u, well-defined on $\mathbb{R} \setminus \{0\}$

$$u'(x) = \frac{-x\lambda}{\sqrt{2\pi}} \int_0^\infty \frac{1}{t^{3/2}} \exp\left(-\frac{x^2}{2}\frac{1}{t} - \lambda t\right) dt, \qquad x \neq 0 \tag{2.9}$$

found by differentiating under the integral. Note that u is not differentiable in 0 as the derivative from the left at 0 approaches $-\sqrt{2\lambda}$ while the derivative from the right approaches $\sqrt{2\lambda}$. Surprisingly, we can rewrite u' from (2.9) in a clever way as

$$u'(x) = \begin{cases} -\lambda e^{-\sqrt{2\lambda}x}, & x > 0, \\ \lambda e^{\sqrt{2\lambda}x}, & x < 0, \end{cases} \tag{2.10}$$

proven below and hence we see that

$$u(x) = \begin{cases} \frac{\sqrt{2\lambda}}{2} e^{-\sqrt{2\lambda}x}, & x > 0 \\ \frac{\sqrt{2\lambda}}{2} e^{\sqrt{2\lambda}x}, & x < 0. \end{cases}$$

We wish for u to be continuous and choose

$$u(x) = \frac{\sqrt{2\lambda}}{2} e^{-\sqrt{2\lambda}|x|}, \qquad x \in \mathbb{R}. \tag{2.11}$$

To prove (2.10), we quote [KT75] p. 363 stating that for $x < 0$ and $\xi \geq 0$

$$1_{(0,\infty)}(t) \frac{-x}{\sqrt{2\pi t^3}} \exp\left(-\xi x - \frac{1}{2}\xi^2 t - \frac{x^2}{2t}\right)$$

is a probability density. In particular it integrates to 1 and by rearranging and substituting $\xi = \sqrt{2\lambda} > 0$ we obtain

$$e^{\sqrt{2\lambda}x} = \frac{-x}{\sqrt{2\pi}} \int_0^\infty \frac{1}{t^{3/2}} \exp\left(-\lambda t - \frac{x^2}{2t}\right) dt, \qquad x < 0,$$

proving (2.10) for $x < 0$. In the case $x > 0$ the identity holds for $-x < 0$ and thus simply substitute x with $-x$ to obtain the desired. In particular using (2.8) and (2.11), we have along the way proved the useful identity

$$\int_0^\infty \frac{1}{\sqrt{2\pi t}} \exp\left(-x^2 \frac{1}{2t} - \lambda t\right) dt = \frac{1}{\sqrt{2\lambda}} e^{-\sqrt{2\lambda}|x|}, \quad x \in \mathbb{R} \setminus \{0\}, \lambda > 0. \tag{2.12}$$

Another way of finding the stationary distribution is to use the infinitesimal generator A from (2.4) here becoming

$$Af(x) = \frac{1}{2} f''(x) + \lambda(f(0) - f(x)), \qquad f \in C_0^2(\mathbb{R}). \tag{2.13}$$

We are interested in finding an initial stationary distribution ν for $\mathbf{X}$ and such ν must satisfy

$$\nu(Af) = 0 \tag{2.14}$$

for $f \in L(\nu)$. As $C_0^2(\mathbb{R})$ is dense in $L(\nu)$ it is a determining class and thus any probability is uniquely determined from the behavior with respect to such f's. In the following we assume that ν has continuous density u with respect to the Lebesgue-measure. One finds

$$\nu(Af) = \int \left[\frac{1}{2} f''(x) + \lambda(f(0) - f(x))\right] u(x)\, dx = \int \left[\frac{1}{2} f'' - \lambda f\right] u\, dx + \lambda f(0). \tag{2.15}$$

As we have done earlier using the space of test functions integration by parts is applied to move the derivatives from f to u and as this only has to be done twice it suffices for f to be C_0^2. While these considerations at first seem innocent a subtle point has been omitted. We have to require that u' and u'' exist at least on a set with probability 1. This assumption is not of pathological nature as should we succeed in our endeavor of finding the same u as in (2.11) then indeed u is not differentiable in 0. Nevertheless, $u'(0^-)$ and $u'(0^+)$ are both well-defined.[1]

Using ordinary analysis, care has to be taken at $x = 0$ and thus the integral is split around 0. We first consider

$$\begin{aligned}\int_{0^+}^\infty f'' u\, dx &= [f'u]_{0^+}^\infty - \int_{0^+}^\infty f'u'\, dx = -f'(0)u(0^+) - \int_{0^+}^\infty f'u'\, dx \\ &= -f'(0)u(0^+) + f(0)u'(0^+) + \int_{0^+}^\infty f u''\, dx.\end{aligned}$$

Analogously, we find

$$\int_{-\infty}^{0^-} f'' u\, dx = \int_{-\infty}^{0^-} f u''\, dt + f'(0)u(0^-) - f(0)u'(0^-).$$

[1] Strictly speaking for an element u in $L(\nu)$ it is also meaningless to talk about $u(0)$, but as we require u to be continuous this uniquely identifies a member of the equivalence class.

Inserting these expressions into (2.15), we get the intermediate result for $f \in C_0^2(\mathbb{R})$

$$\begin{aligned}\nu(Af) = &\int_{0^+}^{\infty} (u''/2 - \lambda u) f \, dt - f'(0)u(0^+)/2 + f(0)u'(0^+)/2 \\ &+ \int_{-\infty}^{0^-} (u''/2 - \lambda u) f \, dt + f'(0)u(0^-)/2 - f(0)u'(0^-)/2 + \lambda f(0) = 0.\end{aligned} \tag{2.16}$$

Restricting our attention to f's with support on $[0, \infty)$ it is seen that u must satisfy $u'' - 2\lambda u \equiv 0$ on $(0, \infty)$ as such f's are dense in $C^2(0, \infty)$. Analogously on $(-\infty, 0)$. Thus

$$u(x) = \begin{cases} A_1 e^{\sqrt{2\lambda}x} + B_1 e^{-\sqrt{2\lambda}x}, & x < 0, \\ A_2 e^{\sqrt{2\lambda}x} + B_2 e^{-\sqrt{2\lambda}x}, & x > 0. \end{cases} \tag{2.17}$$

As u defines a finite measure on $\mathbb{R}$ it follows that $B_1 = A_2 = 0$. In order to determine A_1 and B_2 we insert the values for u into (2.16) and remember that with this particular u the integral of both integrand equals 0. Therefore

$$-f'(0)B_2 - f(0)B_2\sqrt{2\lambda} + f'(0)A_1 - f(0)A_1\sqrt{2\lambda} + 2\lambda f(0) = 0.$$

Choosing f such that $f(0) = 0$ this yields $A_1 = B_2$, while an f such that $f'(0) = 0$ yields

$$-f(0)B_2\sqrt{2\lambda} - f(0)A_1\sqrt{2\lambda} + 2\lambda f(0) = 0.$$

Summing up $A_1 = B_2 = \sqrt{2\lambda}/2$ such that

$$u(x) = \frac{\sqrt{2\lambda}}{2} e^{-\sqrt{2\lambda}|x|}, \qquad x \in \mathbb{R}.$$

This could also be found using (2.17) directly, that u is continuous to give $A_1 = B_2$ and finally that u has to integrate to 1.

Checking the solution To check if the solution actually fulfills the differential equation and the boundary condition one might be tempted—since u is continuous—to check that

$$\int_{\mathbb{R}} (1/2u'' - \lambda u) f dx = \int_{\mathbb{R}\setminus\{0\}} (1/2u'' - \lambda u) f dx = -\lambda f(0), \quad f \in C^2(\mathbb{R}). \tag{2.18}$$

Alas, the second integral in the above equals 0! This is hardly surprising as u on $(-\infty, 0)$ and $(0, \infty)$, respectively, was chosen to be a function for which $1/2u'' - \lambda u \equiv 0$ and thus the integrand is zero. So where is the flaw in our argument? While u is continuous u'' is not. In fact, it is not even a function, but rather a distribution with atoms in 0 and thus $\{0\}$ is not a null-set for the integrand.[2]

We wish to find $\partial^2 u$ in a distributional sense, see Section A.1. First, we find ∂u and as u is continuous and piecewise differentiable on $(-\infty, 0)$ and $(0, \infty)$ its derivative in distribution sense $\partial u \in L_1(\mathbb{R})$ is given by

$$\partial u(x) = \left(\frac{\sqrt{2\lambda}}{2} e^{-\sqrt{2\lambda}|x|} \right)' = \begin{cases} \lambda e^{\sqrt{2\lambda}x} & \text{on } (-\infty, 0) \\ -\lambda e^{-\sqrt{2\lambda}x} & \text{on } (0, \infty) \end{cases} = \operatorname{sgn}(x)(-\lambda) e^{-\sqrt{2\lambda}|x|}.$$

[2] The measure-theoretic statement is that u'' does not have a density with respect to the Lebesgue measure and thus it is not automatic that $\{0\}$ is a null-set.

Albeit exotic at first glance it fits our expectations and the regular chain rule as $\partial|x| = \operatorname{sgn}(x)$. Nevertheless, it is still a distribution function with no atoms, but things turn a bit more involved when we wish to find $\partial^2 u$ and we need to invoke a bit more of distribution theory. Let $\varphi \in C_0^\infty(\mathbb{R})$ be a test function. As the inner product operator (A.1) is linear, we can rest-assured omit the constant factor $-\lambda$ from our calculations at first and using (A.2) it is seen that

$$\begin{aligned}
&\left\langle \partial \operatorname{sgn}(x) e^{-\sqrt{2\lambda}|x|}, \varphi \right\rangle = -\left\langle \operatorname{sgn}(x) e^{-\sqrt{2\lambda}|x|}, \partial\varphi \right\rangle \\
&= \int_{-\infty}^{0} e^{\sqrt{2\lambda}x} \varphi' \, dx - \int_{0}^{\infty} e^{-\sqrt{2\lambda}x} \varphi' \, dx \\
&= [e^{\sqrt{2\lambda}x}\varphi]_{-\infty}^{0} - \int_{-\infty}^{0} \sqrt{2\lambda} e^{\sqrt{2\lambda}x} \varphi \, dx - [e^{-\sqrt{2\lambda}x}\varphi]_{0}^{\infty} - \int_{0}^{\infty} \sqrt{2\lambda} e^{-\sqrt{2\lambda}x} \varphi \, dx \\
&= \varphi(0) - \left\langle 1_{(-\infty,0)}(x) \sqrt{2\lambda} e^{\sqrt{2\lambda}x}, \varphi \right\rangle + \varphi(0) - \left\langle 1_{(0,\infty)}(x) \sqrt{2\lambda} e^{-\sqrt{2\lambda}x}, \varphi \right\rangle \\
&= \left\langle 2\delta_0 - \sqrt{2\lambda} e^{-\sqrt{2\lambda}|x|}, \varphi \right\rangle,
\end{aligned}$$

Having omitted a factor $-\lambda$ it follows that in a distributional sense

$$\partial^2 u(x) = -2\lambda\delta_0 + \lambda\sqrt{2\lambda} e^{-\sqrt{2\lambda}|x|}$$

with δ_0 the Dirac delta-distribution. Success is close and we see

$$1/2 \partial^2 u(x) - \lambda u(x) = -\lambda\delta_0$$

thus satisfying $\int -\lambda\delta_0 f(x)\, dx = \lambda f(0)$ as desired.

2.4.2. General Brownian motion

Keeping degenerate Y_i's, we can treat the case of a general Brownian motion $X_t^{(i)}$ with drift θ and variance σ^2 in the following theorem.

Theorem 2.4. *Consider the jump-diffusion with jump targets model where $Y_i \equiv \varepsilon_0$, the jumps are Poisson with intensity λ, and where the diffusion parts $\mathbf{X}^{(i)}$ are general Brownian with drift θ and variance σ^2 for all $i \in \{1, 2 \ldots\}$. A stationary initial distribution exists and has density*

$$u(x) = \frac{\lambda}{\sqrt{\theta^2 + 2\sigma^2\lambda}} \exp\left(-\frac{\sqrt{\theta^2 + 2\sigma^2\lambda}|x| - \theta x}{\sigma^2} \right). \tag{2.19}$$

In particular for a standard Brownian motion it is the Laplace distribution with density

$$u(x) = \frac{\sqrt{2\lambda}}{2} \exp\left(-\sqrt{2\lambda}|x| \right). \tag{2.20}$$

Proof. The density of a general Brownian motion at time t is given by

$$p_t(0, x) = \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left(-\frac{(x - \theta t)^2}{2\sigma^2 t} \right).$$

Since the jumps are Poisson with intensity λ the stationary renewal process distribution is $f_\rho(t) = \lambda \exp(-\lambda t)$ for $t > 0$. Using (2.7) it is seen that

$$u(x) = \int_0^\infty dt\, f_\rho(t) p_t(0, x) = \frac{\lambda}{\sigma\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{t}} \exp\left(-\frac{(x - \theta t)^2}{2\sigma^2 t} - \lambda t \right) dt.$$

Rewriting this and using (2.12) gives

$$\begin{aligned} u(x) &= \frac{\lambda \exp(\frac{\theta x}{\sigma^2})}{\sigma} \int_0^\infty \frac{1}{\sqrt{2\pi t}} \exp\left(-\left(\frac{x}{\sigma}\right)^2 \frac{1}{2t} - \left(\frac{\theta^2}{2\sigma^2} + \lambda\right) t\right) dt \\ &= \frac{\lambda \exp(\frac{\theta x}{\sigma^2})}{\sigma} \cdot \frac{1}{\sqrt{\theta^2/\sigma^2 + 2\lambda}} e^{-\sqrt{\theta^2/\sigma^2+2\lambda}\,|\frac{x}{\sigma}|} \\ &= \frac{\lambda}{\sqrt{\theta^2 + 2\sigma^2\lambda}} \exp\left(-\frac{\sqrt{\theta^2 + 2\sigma^2\lambda}|x| - \theta x}{\sigma^2}\right), \qquad x \in \mathbb{R} \setminus \{0\} \end{aligned}$$

since $\sigma > 0$. □

2.5. Brownian motion with arbitrary jump distribution

Maintaining jumps from a Poisson process, this section allows for more general distributions of the Y_i's with density g. As before we begin by considering a standard Brownian motion treated in Theorem 2.5 while the stationary density for the general Brownian motion is found in Theorem 2.6.

2.5.1. Standard Brownian motion

We begin by proving a result very similar to Theorem 2.4.

Theorem 2.5. *Consider the jump-diffusion with jump targets model where Y_i has density g for all $i \in \{0, 1, \ldots\}$, the jumps are Poisson with intensity λ, and where the diffusion parts are standard Brownian. A stationary initial distribution exists and has density*

$$u(x) = \left(g * \frac{\sqrt{2\lambda}}{2} \exp\left(-\sqrt{2\lambda}|z|\right)\right)(x). \tag{2.21}$$

Proof. The strategy is to use the condition on the infinitesimal generator to obtain a differential equation, which is then solved by Fourier transformation.

Since f has compact support the derivatives can be moved freely between f and u merely adjusting the sign appropriately. Using this combined with the fact that g and u both integrate to 1, one finds

$$\begin{aligned} \nu(Af) &= 1/2 \int f'' u dx + \lambda \int \int (f(y) - f(x)) g(y) u(x) dy dx \\ &= 1/2 \int f u'' dx + \lambda \left[\int f g dy - \int f u dx\right] \\ &= 1/2 \int f u'' dx + \lambda \int (g - u) f dx \\ &= \int [u''/2 + \lambda(g - u)] f dx \end{aligned}$$

for $f \in C_0^2(\mathbb{R})$. These calculations are valid only if u' and u'' are defined almost everywhere as one is acting under the integral sign. Hence for the following the analysis to be true the u considered must be C^2 on a set of full measure and integrate to 1.[3] This is checked when a candidate for u has been identified.

[3] The calculations above have not assumed that u is a density. This is, of course, a requirement stated in the boundary condition, but for the analysis to be true thus far it is not needed.

Using the above expression and solving $\nu(Af) = 0$ one obtains

$$u'' - 2\lambda u = -2\lambda g \qquad \text{with boundary condition } u \text{ a density} \tag{2.22}$$

the Fourier transform of which becomes

$$-s^2\hat{u}(s) - 2\lambda\hat{u}(s) = -2\lambda\hat{g}(s)$$

and isolating $\hat{u}$

$$\hat{u}(s) = \sqrt{2\lambda}\hat{g}(s) \cdot \frac{\sqrt{2\lambda}}{s^2 + 2\lambda}.$$

By (A.10) and (A.11) this is the Fourier transform of

$$u(t) = \sqrt{2\lambda}\sqrt{\pi/2}\frac{1}{\sqrt{2\pi}}\left(g * e^{-\sqrt{2\lambda}|z|}\right)(t) = \left(g * \frac{\sqrt{2\lambda}}{2}e^{-\sqrt{2\lambda}|z|}\right)(t), \quad t \in \mathbb{R}. \tag{2.23}$$

To check the requirements for u recall that two probability densities convolved again yields a probability density and as such u is indeed a probability density. It remains to verify that u is twice differentiable. If g is a continuous density then by the smoothing property of convolution u is smooth everywhere with the only possible exception of $x = 0$. If g has atoms each of these leads to a singularity for u as convolution is a linear operation and $(\varepsilon_{z_0} * h)(x) = h(x - z_0)$. Since g is a density it can at most have countably many singularities, which carries over to u. Therefore u is C^2 on a set of full measure as required. □

Note that this theorem does not by itself give any information about the behavior of the point process producing the jumps. It is assumed to be Poisson and the density for the resulting stationary distribution naturally depends on this choice, but one cannot isolate the effects of the point process in the density. We have a priori no chance of concluding anything about either the $X^{(i)}$'s, the Y_i's or the T_i's, but only about their interplay and how this forms a stationary distribution. Therefore Corollary 2.3 is useful.

As discussed in Section A.2 convolution is a smoothing operation and it is tempting to conclude that as $\exp(|x|)$ is infinitely smooth everywhere except at $x = 0$ the same must hold for u. Nevertheless, as we saw in the proof above this is only true if g is a function whereas if g is a measure with atoms the argument fails. For instance for $g \equiv (\varepsilon_{-1} + \varepsilon_1)/2$ both -1 and 1 are singularities of u as illustrated in Figure 2.1.

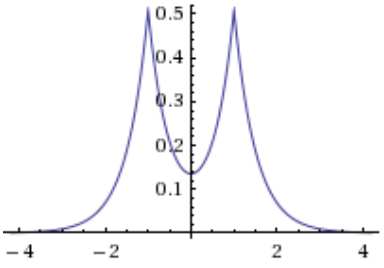

Figure 2.1.: Plot of $u = 1/2[\exp(-2|t-1|) + \exp(-2|t+1|)]$ with two singularities at -1 and 1, respectively, being the resulting distribution for $g \equiv (\varepsilon_{-1} + \varepsilon_1)/2$ with $\lambda = 2$.

Having done the work in the previous section it is easy to verify that u is a solution to (2.22). With φ a test function

$$\left\langle \partial^2 u, \varphi \right\rangle = \left\langle \left(g * \partial^2\sqrt{2\lambda}/2e^{\sqrt{2\lambda}|x|}\right)(t), \varphi \right\rangle = \left\langle \left(g * (-2\lambda\delta_0 + \lambda\sqrt{2\lambda}e^{-\sqrt{2\lambda}|x|})\right)(t), \varphi \right\rangle.$$

Finally, recalling that the density of the sum of two independent random variables is the convolution of their densities, u of Theorem 2.5 is in fact the density of $X_t + Y_i$, where $\mathbf{X}$ is the Brownian motion being reset to 0 at every jump and initialized with its stationary distribution, while Y_i continues to denote the jumps. The intuition behind this is that since $\mathbf{X}$ and Y_i are independent any jump can be seen as a two-step experiment where a point process gives the jump time while a second independent experiment yielding the jump target Y_i is undertaken.

Note that (2.21) with $g \equiv \varepsilon_0$ reduces to (2.20) and as such one could easily have used the Fourier transform technique from the proof of Theorem 2.5 to treat the special case presented in Theorem 2.4.

2.5.2. General Brownian motion

The most general version of the theorems above for a general Brownian motion, Poisson jumps, and arbitrary jump target distribution is contained in the following theorem. The proof contains no surprises and is a straight-forward generalization of the proof of Theorem 2.5. Since no additional insight is offered a sketch highlighting the differences is presented.

Theorem 2.6. *Consider the jump-diffusion with jump targets model where Y_i has density g for all $i \in \{0, 1, \ldots\}$, the jumps are Poisson with intensity λ, and where the diffusion parts are general Brownian with drift θ and variance σ. A stationary initial distribution exists and has density*

$$u \equiv \left(g * \frac{\lambda}{\sqrt{\theta^2 + 2\sigma^2\lambda}} \exp\left[-\frac{\sqrt{\theta^2 + 2\sigma^2\lambda}|x| - \theta x}{\sigma^2} \right] \right). \tag{2.24}$$

Proof sketch. The ODE to be solved is

$$\sigma^2 u'' - 2\theta u' - 2\lambda u = -2\lambda g$$

for which the Fourier transform becomes

$$\hat{u}(s) = 2\lambda \hat{g}(s) \cdot \frac{1}{s^2\sigma^2 + 2\theta i s + 2\lambda}.$$

Using (A.12) it is seen that

$$\mathcal{F}\left(\frac{\lambda}{\sqrt{\theta^2 + 2\sigma^2\lambda}} \exp\left[-\frac{\sqrt{\theta^2 + 2\sigma^2\lambda}|x| - \theta x}{\sigma^2} \right] \right)(s) = \sqrt{\frac{2}{\pi}} \cdot \frac{\lambda}{s^2 + 2\lambda + 2\theta i s}.$$

Applying the convolution theorem for Fourier transforms (A.10) yields the desired. □

2.6. Stationarity for general diffusion pieces and non-constant jump intensity

So far the analysis has assumed constant jump-intensity and the Brownian motion, but we shall now consider the situation for general $\lambda(x)$ and general diffusions. Integrals such as (2.7) are inaccessible with general $\lambda(x)$. Instead obtaining a differential equation for the stationary initial distribution from solving $\nu(Af) = 0$ is the path we choose to generalize.

So far these differential equations have been sufficiently simple to be solved, but this is, unsurprisingly, not always the case. When $\lambda(x)$ is non-constant the task is further complicated as an integral factor is introduced into the equation turning it into a integro-differential equation. We present the result in the following theorem.

Theorem 2.7 (Stationarity criterion)**.** *Consider a jump-diffusion with jump targets model where Y_i has density g for all $i \in \{0, 1, \ldots\}$, the jumps have x-dependent intensity $\lambda(x)$, and where the diffusion parts are given by (2.1)*

$$dX_t^{(i)} = b(X_t^{(i)})dt + \sigma(X_t^{(i)})dB_t, \qquad X_0^{(i)} = Y_{i-1}.$$

The stationary density u, if it exists, must satisfy

$$\begin{aligned} \frac{\sigma^2(x)}{2}u''(x) + \Big[(\sigma^2)'(x) - b(x)\Big]u'(x) + \Big[\frac{1}{2}(\sigma^2)''(x) - b'(x) - \lambda(x)\Big]u(x) \\ = -g(x)\int \lambda(y)u(y)dy \end{aligned} \tag{2.25}$$

with boundary condition 'u a density.'

Remark By Theorem 2.1 u exists if $\mu = \mathrm{E}_G[T_1] < \infty$.

Proof. Assume u exists. The necessary and sufficient condition $\nu(Af) = 0$ from (2.14) remains true and familiar calculations show that

$$\begin{aligned} \nu(Af) &= \int bf'u + \sigma^2/2f''udx + \int \lambda(x)u(x)\int [f(y) - f(x)]g(y)dydx \\ &= \int [-(bu)' + (\sigma^2/2u)'']fdx + \int \lambda udx \int fgdy - \int \lambda ufdx \\ &= \int \Big[- (bu)' + (\sigma^2/2u)'' - \lambda u + g\int \lambda udy\Big]fdx \end{aligned}$$

for $f \in C_0^2(\mathbb{R})$. Solving $\nu(Af) = 0$ is thus equivalent to the following integro-differential equation in u

$$-(bu)'(x) + \frac{1}{2}(\sigma^2 u)''(x) - \lambda(x)u(x) + g(x)\int \lambda(y)u(y)dy = 0,$$

which reduces to (2.25) as desired. □

Despite its monstrous appearance, (2.25) deserves a right of existence. If one somehow can produce a candidate for a stationary distribution—by guess, divine inspiration, et cetera—it is a very explicit condition to check. In particular for constant jump intensities it is nearly mechanical to check.

The above gives a description of the stationary distribution using the generator and the fact that $\nu(Af) = 0$ is a sufficient and necessary criterion for ν to be stationary. However, we began the chapter by using a renewal theory approach in short using the transition probabilities of the Brownian motion are known. This allowed us to find the density of B_{T_1} where T_1 is an exponential waiting time. Generally this argument gives:

Theorem 2.8. *Let the jump intensity be constant λ, the diffusion have transition densities $(p_t(y, x))_t$, and the jump target Y_i have density g for all $i \in \{0, 1, \ldots\}$. Then the stationary density u is given by*

$$u \equiv g * \int_0^\infty dt\, \lambda e^{-\lambda t} p_t(0, x). \tag{2.26}$$

Proof. If $Y_i \equiv \varepsilon_0$ it follows that

$$u \equiv \int_0^\infty dt\, \lambda e^{-\lambda t} p_t(0,x).$$

Consider general Y_i with density. As Y_i is independent of the diffusion pieces as well as the renewal process the resulting stationary distribution is simply the convolution of the densities yielding (2.26). □

The significance of Theorem 2.8 lies in the fact that for an arbitrary diffusion $\mathbf{W}$ the density of W_{T_1} where $T_1 \sim \exp(\lambda)$ and independent of $\mathbf{W}$, that is

$$\int_0^\infty dt\, \lambda e^{-\lambda t} p_t(0,z),$$

is of general interest. The standard and general Brownian motion at exponential stopping has already been treated in the above. For the GBM see (1.41) and for the OU see (1.47), while there is no expression familiar to the author for the CIR-process. The very (!) lengthy expressions of the exponential stopping density for the GBM- and OU-processes explains why we have only considered the Brownian motion.

2.7. Stationary diffusions

Some diffusions such as the Brownian motion do not themselves have a stationary distribution, while others such as the Cox-Ingersoll-Ross process (CIR) under certain conditions do, see Theorem 1.8. We describe the general idea before turning to the specifics of the CIR. Let $(Z_t)_{t\geq 0}$ be a diffusion with a stationary initial distribution ν_D. The idea to obtain a stationary distribution for the jump-diffusion is now very simple: With Poisson jumps, using ν_D as the distribution of all the Y_i's, we would expect that the jump-diffusion is again stationary with $\nu = \nu_D$.

First note that for this setup it in general holds that (2.25) reduces to

$$\frac{\sigma^2(x)}{2}u''(x) + \Big[(\sigma^2)'(x) - b(x)\Big]u'(x) + \Big[\frac{1}{2}(\sigma^2)''(x) - b'(x)\Big]u(x) = 0 \qquad (2.27)$$

Recall from Section 1.3.4.4 that the CIR process is given as the solution to the SDE

$$dZ_t = a + bZ_t dt + \sigma\sqrt{Z_t}dB_t, \qquad Z_0 = z_0 > 0$$

and that a stationary distribution exists if and only if (see (1.50))

$$\frac{2a}{\sigma^2} \geq 1 \quad \text{and} \quad b < 0.$$

To simplify the computation, we put $2a = \sigma^2$ and $b = -1$ such that the stationary density becomes exponential (see (1.51))

$$u(x) = \frac{2}{\sigma^2}\exp(-2/\sigma^2 x) = \frac{1}{a}\exp(-x/a), \qquad a > 0.$$

The left-hand side of (2.27) becomes

$$\exp(-x/a)\Big[\frac{\sigma^2 x}{2}\frac{1}{a^3} - (\sigma^2 - a - bx)\frac{1}{a^2} - \frac{b}{a}\Big] = \exp(-x/a)\Big[\frac{x}{a^2} - \frac{a+x}{a^2} + \frac{1}{a}\Big] = 0$$

as desired.

This example motivates the following theorem generalizing the above result and offering somewhat of a partial converse statement.

Theorem 2.9. *Let the diffusion pieces be arbitrary* $\mathbf{X}^{(i)}$ *for which a stationary distribution*ν_D *exists.*

Consider the jump-diffusion with jump targets model where Y_i *has distribution* ν_D *and the jumps are Poisson. The stationary distribution for the jump-diffusion exists and is* ν_D*.*

Conversely, if the stationary distribution for the jump-diffusion ν *exists and* Y_i *has distribution* ν *for all* $i \in \{0, 1, \ldots\}$ *then* $\nu \equiv \nu^D$*.*

Proof. With Poisson with intensity λ the stationary distribution for the jump-diffusion always exists. Recall from (2.4) that $Af(x) = A^d f(x) + \lambda \int f(y) - f(x)\, g(dy)$ such that

$$\nu(Af) = \int Afu\, dx = \int A^d fu\, dx + \lambda \int (g - u) f\, dx, \qquad f \in C^2(\mathbb{R}). \tag{2.28}$$

From Theorem 1.8 it follows that if u is the stationary density of the diffusion then $\int A^d fu\, dx = 0$. By choosing $g \equiv u$ it follows that $\nu(Af) = 0$ showing that u is the stationary for the jump-diffusion too.

For the second part of the theorem rewrite (2.28) as

$$\int A^d fu\, dx = \nu(Af) - \lambda \int (g - u) f\, dx.$$

By assumption $\nu(Af) = 0$ and $g \equiv u$ for all $f \in C^2(\mathbb{R})$ implying $A^d fu \equiv 0$ such that u is the density of the stationary distribution for the diffusion by Theorem 1.8. □

The above theorem requires that the jump intensity is constant and to round off this section we discuss in a colloquial setting what happens when this is not the case. A constant jump intensity means that the time spent in the state solely determines the probability for the next jump to occur. On the other hand, letting the jump intensity depend on the position of $\mathbf{X}$—say larger x leads to larger intensities—introduces some askewness in the stationary distribution for the jump-diffusion and pulls it closer towards the distribution of Y_i for such x leading to large intensities.

Metaphorically speaking, it is possible to discriminate certain regions of the diffusion by attaching high intensities to these parts such that a jump is more likely to occur and the time spent in that region thereby decreases. A constant intensity can be thought of as non-discriminating and is much easier to grasp both intuitively as we can rely on our extensive knowledge of standard diffusions as well as mathematically.

This description is a bit simplified and for instance it is very well conceivable that regions with low intensities are only reached with a very small probability. Think for instance of a target distribution resetting to 0 at every jump and where the diffusion is a Brownian motion with a large positive drift. Attaching very high intensity to the immediate region of a for instance using $\lambda(x) = \exp[(x - a)/b]$ for b minute and positive, results in a very low probability to go beyond a. The stationary probability being in a small neighborhood around 0 is small. Such an intensity can be used to model a threshold a.

CHAPTER 3

Jump-diffusion models using jump sizes

This chapter builds a model of jump-diffusions closely resembling the model of Chapter 2. The model developed here and referred to as a *jump-diffusion with jump sizes* specifies at times of jumps the change in position rather than the post-jump target.

As it turns out, this poses quite a different challenge mathematically and to the machinery employed to address this model. The main purpose of this chapter continuous to be to investigate the stationary distribution of the jump-diffusion. We begin by establishing a simple criterion for the existence of such where after we analyze various models of interest. Unfortunately, explicitness is much harder to obtain in this model.

3.1. The model and stationary distributions

With the setup and notation known from Chapter 2 except now letting the Y_i's describe jump sizes rather than targets, we still assume the Y_i's to be iid with distribution G and if a density exists g. We require that $\mathrm{E}[Y_i] < \infty$ and $Y_i \geq 0$ a.s., both requirements not previously needed in the jump target model.

The jump-diffusion is denoted by $\mathbf{X}$ while $\mathbf{X}^{(i)}$ are the diffusions-pieces and at jump times $T_1, T_2, \ldots$, we find

$$X_{T_i} = X_{T_i}^{(i)} + Y_i.$$

Let the distribution of $\mathbf{X}_0^{(1)}$ be H with density κ when it exists. If we can somehow make sure that $X_{T_1} + Y_1$—the starting point for $\mathbf{X}^{(2)}$—has distribution H, too, then the $\mathbf{X}^{(i)}$'s are mutual probabilistic replicas. Using the renewal approach $\mathbf{X}$ then becomes a regenerative process for which a unique stationary distribution is known to exist. If in addition X_{T_i} has distribution L, we get the following relation

$$H = (L \star G) \tag{3.1}$$

There is one catch though: The argument relies on the fact that the process is being renewed and thus we must assume that $\mathrm{E}_\kappa[T_1] < \infty$.

In the search of such H let ν_n be the distribution of $X_0^{(n+1)}$ with $n \geq 0$ and put $X^{(0)} \equiv 0$ such that

- ν_0: The distribution of Y_1,
- ν_n, $n \geq 1$: The distribution of $X_{T_1}^{(1)} + Y_1$ when $X_0^{(1)}$ has distribution ν_{n-1}.

where it is used that the $X^{(i)}$'s are probabilistic replicas if the Y_i's are iid. Should $\nu_\infty := \lim_{n\to\infty} \nu_n$ exist this qualifies as H. The densities of ν_n, when they exist, are denoted v_n and ν_∞ then has density v_∞. We get the following recurrence relation

$$\nu_{n+1}(f) = \mathrm{E}_{\nu_n}[f(X_{T_1}^{(1)} + Y_1)] = \iint \mathrm{E}_x[f(X_{T_1}^{(1)} + y)]\, d\nu_n(x)\, dG(y), \quad n \geq 1 \tag{3.2}$$

where the auxiliary notation of $h_f(x, y)$ is introduced

$$\begin{aligned} h_f(x,y) &:= \mathrm{E}_x[f(X_{T_1}^{(1)} + y)] \\ &= \mathrm{E}_x\left[\int_0^\infty dt\, \lambda(X_t^{(1)}) e^{-\int_0^t \lambda(X_s^{(1)})ds} f(X_t^{(1)} + y)\right]. \end{aligned} \tag{3.3}$$

Recall that $\mathrm{E}_{v_n}[X_t^{(1)}]$ denotes the expectation of $X_t^{(1)}$ when started according to ν_n. Both the notion of $h_f(x, y)$ and $\mathrm{E}_x[f(X_{T_1}^{(1)} + y)]$ are used throughout.

The following definition combined with a simple result turns out to be useful when tackling the problem of ν_∞'s existence.

Definition 3.1 (Stochastic ordering)**.** Let η_1 and η_2 be two measures on $(\overline{\mathbb{R}}, \overline{\mathcal{B}})$. We say that η_2 dominates η_1 stochastically if

$$\nu_1([x, \infty]) \leq \nu_2([x, \infty]) \qquad \text{for all } x \in \overline{\mathbb{R}} \tag{3.4}$$

and write $\eta_1 \preceq \eta_2$. If W_1 and W_2 are stochastic variables with distribution η_1 and η_2, we write $W_1 \preceq W_2$.

Formula (3.4) is equivalent to

$$\eta_1(f) = \mathrm{E}[f(W_1)] \leq \mathrm{E}[f(W_2)] = \eta_2(f) \tag{3.5}$$

for f bounded and non-decreasing if $W_1 \preceq W_2$.

We merely quote the result that if $\nu_1 \preceq \nu_2 \preceq \ldots$ then a weak limit $\nu_\infty := \lim_{n\to\infty} \nu_n$ on $(-\infty, \infty]$ exists. Note that we cannot be sure that ν_∞ stays away from ∞. This is certainly not the case in general as seen by considering $\nu_n := \varepsilon_n$.

Relaying renewal theory and the preceding analysis we formulate the following existence theorem for the jump size model. The idea is to realize how certain jump size models can be reformulated as a sort of jump target models.

Theorem 3.1. *Consider the jump-diffusion model with jump sizes. Let the limit distribution $\nu_\infty := \lim_{n\to\infty} \nu_n$ exists, be a probability on $(\mathbb{R}, \mathcal{B})$,*

$$\mu = \mathrm{E}_{\nu_\infty}[T_1] < \infty, \tag{3.6}$$

and all $\mathbf{X}_0^{(i)} \overset{\mathcal{D}}{=} Y_i$. Then $\mathbf{X}$ has a unique stationary distribution.

Proof. Theorem 2.1 gives sufficient conditions for the existence of a stationary distribution exists in a jump target model. It is to be shown how a jump size model under the above assumptions can be reformulated such that Theorem 2.1 can be applied.

By assumption ν_∞ on $\mathbb{R}$ exists and can be used as the jump target distribution G. Since all diffusion parts are started according to ν_∞ and it is assumed $\mathrm{E}[T_1] < \infty$, Theorem 2.1 yields the existence of a stationary density for this particular jump size model. □

It is critically important to note that ν_∞ is not the stationary distribution itself. Rather, it is simply the distribution with which all the $\mathbf{X}^{(i)}$ must be started such that $\mathbf{X}$ has a stationary distribution.

Also note that ν_∞ depends on the jump intensity, but we usually suppress this in the notation. However, when required we denote for $\lambda(x)$ the resulting measure $\nu_\infty^{\lambda(x)}$.

3.2. Deterministic pieces

To get started we must make some simplifying assumptions. Unsurprisingly, it is a lot simpler to consider jumps from a Poisson process, but to begin with we also use deterministic pieces of the jump-"diffusion." These are denoted by $\varphi_s(x)$, where x is the starting point and s the time passing, see Section 1.5 in particular (1.61). Thus $\mathbf{X}$ is a piecewise deterministic process.

With the $\mathbf{X}^{(i)}$'s deterministic, (3.3) becomes particular easy to handle as for instance

$$\mathrm{E}_x[f(X_s^{(1)}] = \varphi_x(s).$$

Thus if $\mathbf{X}^{(1)}$ is started with distribution H and we assume κ exists it holds that

$$\begin{aligned}\mathrm{E}_\kappa[f(X_{T_1}^{(1)})] &= \int \kappa(x) \int_0^\infty \lambda e^{-\lambda s} f(\varphi_s(x))\, ds\, dx \\ &= \int_0^\infty \lambda e^{-\lambda s} \int \kappa(\varphi_s^{-1}(z)) f(z) \frac{1}{|\varphi_s^{-1}(z)'|}\, dz\, ds\end{aligned} \tag{3.7}$$

if $\varphi_s(x)$ is a diffeomorphism. η being the density of X_{T_1} must satisfy $\mathrm{E}_\kappa[f(X_{T_1}^{(1)})] = \int f(z)\eta(z)\, dz$ and changing the order of integration in (3.7) gives

$$\eta(z) = \int_0^\infty \lambda e^{-\lambda s} \frac{1}{|\varphi_s^{-1}(z)'|} \kappa(\varphi_s^{-1}(z))\, ds.$$

So far we have only stated that the pieces are deterministic, but not their exact behavior. The most simple example might well be the step process given by $\varphi_s(x) = x$; does this have a chance of converging? Since all jumps are positive and $\varphi_s(x)$ does not decrease we eventually end up with $\nu_\infty = \varepsilon_\infty$ thus diverging. Instead we need something that shoots towards 0 between jumps. We try our luck with $\varphi_s(x) = xe^{-\alpha s}$ with $\alpha, x > 0$ and in this particular case find

$$\eta(z) = \int_0^\infty \lambda e^{-(\lambda+\alpha)s} \kappa(ze^{\alpha s})\, ds$$

which using the substitution $w = ze^{\alpha s}$ can be written as

$$\eta(z) = \frac{\lambda}{\alpha} z^{\lambda/\alpha - 1} \int_z^\infty w^{-\lambda/\alpha} \kappa(w)\, dw. \tag{3.8}$$

The latter form is preferred as it depends on λ and α solely through their ratio, which is intuitively appealing.

It turns out that $\varphi_s(x) = xe^{-\alpha s}$ is a prudent choice and it is kept throughout this section, while other deterministic pieces are briefly considered at the end. So far in our discussion we have left the jump distribution out of the pictures, but of course this has to be specified in order to determine ν_∞ (if it exists).

With Poisson jumps we move on to first consider degenerate jumps and the moving on to general $Y_i \geq 0$'s a.s. with $\mathrm{E}[Y_i] < \infty$.

3.2.1. Special case with degenerate jumps

Let Y_1 have degenerate distribution ε_a for some $a > 0$. In this setup we first want to show that ν_n converges for $n \to \infty$ and then show that ν_∞ has no mass at ∞. The first is shown by means of induction whereas the latter is shown by proving $\nu_\infty(id) < \infty$.

The induction start is to show that $\nu_1 \preceq \nu_2$. This is easy to realize as $\nu_1 = \varepsilon_a$ and $X_{T_1}^{(1)} = \varphi_{T_1}^{(i)}(x) = xe^{-\alpha T_1} > 0$ implying that $X_{T_1}^{(1)} + Y_1$ is always greater than a.

Concerning the induction step assume $\nu_{n-1} \preceq \nu_n$. As a special case of (3.2), we have

$$\nu_{n+1}(f) = \mathrm{E}_{\nu_n}[f(X_{T_1}^{(1)} + a)] = \int \mathrm{E}_x[f(X_{T_1}^{(1)} + a)]\, d\nu_n(x), \qquad n \geq 1 \tag{3.9}$$

where $h_f(x,a) = \mathrm{E}_x[f(X_{T_1}^{(1)} + a)] = \int_0^\infty \lambda e^{-\lambda s} f(xe^{-\alpha s} + a)\, ds$. For any f non-decreasing $x \mapsto h_f(x,a)$ becomes non-decreasing as well and using $\nu_{n-1} \preceq \nu_n$ it is seen that

$$\int h_f(x,a)\, d\nu_{n-1} = \nu_{n-1}(h_f(x,a)) \leq \nu_n(h_f(x,a)) = \int h_f(x,a)\, d\nu_n$$

equivalent to

$$\nu_n(f) \leq \nu_{n+1}(f)$$

which is again equivalent to $\nu_n \preceq \nu_{n+1}$. As $\preceq$ is a transitive relation we have therefore proved $\nu_1 \preceq \nu_2 \preceq \ldots$ implying that ν_∞ exists.

We wish to investigate the behavior of $\nu_\infty \equiv \varepsilon_\infty$ at ∞. The strategy is to show that the expectation of ν_∞ is finite limit implying that $\nu_\infty(\{\infty\}) = 0$. We thus want to find conditions under which

$$\nu_\infty(id) = \int x\, d\nu_\infty < \infty.$$

Using (3.9) with $f \equiv id$ it follows that

$$\nu_{n+1}(id) = \int a + x\frac{\lambda}{\alpha+\lambda}\, d\nu_n(x) = a + \frac{\lambda}{\alpha+\lambda}\nu_n(id), \qquad n \geq 1.$$

This describes a geometric series and inductively we find

$$\nu_n(id) = a\sum_{i=0}^{n}\left(\frac{\lambda}{\lambda+\alpha}\right)^i \tag{3.10}$$

being finite if and only if $\alpha > 0$ in which case

$$\nu_\infty(id) = a\left(1 + \frac{\lambda}{\alpha}\right). \tag{3.11}$$

It seems intuitive that the expectation value depends on the ratio λ/α. When increased we get more jumps each adding a, which should lead to a higher expectation value. That we must require $\alpha > 0$ is soothing as otherwise the deterministic pieces would increase over time pushing up towards ∞.

	$X_{T_1}^{(1)} > 0$	$Y_i \geq 0$
Induction start	✓	–
Induction step	–	✓

Table 3.1.: In the case were the $X_{T_1}^{(i)}$ are deterministic and Y is general the table shows which part of the induction needs which assumptions. Non-negativity of both are thus needed for the induction to work.

3.2.2. General jump size distribution

Keeping deterministic $X_{T_1}^{(i)} = xe^{-\alpha T_1}$, we wish to treat general Y_i rather than degenerate. The ambition is still to show that ν_∞ exists and has finite mean. In order for the analysis to work some assumptions have to be made. These are noted along the way and summarized in Table 3.1.

To show the monotone stochastic ordering we once more use induction. The induction start is obvious as before as long as the deterministic pieces are positive. The induction step is not much more difficult as it still holds that non-decreasing f's lead to non-decreasing $x \mapsto h_f(x, y)$'s. For such a h_f the induction assumption implies

$$\int h_f(x, y) d\nu_{n-1}(x) \leq \int h_f(x, y) d\nu_n(x) \qquad \text{for all } y.$$

Integrating on both sides with respect to the distribution of Y and using (3.2) yields the transition probability and we get the desired inequality $\nu_n(f) \leq \nu_{n+1}(f)$ for any non-decreasing f.

Note whereas the induction step holds for all y the induction start only holds for $y \geq 0$ and thus so does the result.

To compute the expectation value of ν_{n+1} we must consider

$$h_{id}(x, y) = y + \frac{\lambda}{\lambda + \alpha} x$$

from which it follows that

$$\nu_{n+1}(id) = \int_0^\infty \int y + \frac{\lambda}{\lambda + \alpha} x \, d\nu_n(x) \, dG(y) = \mathrm{E}[Y] + \frac{\lambda}{\lambda + \alpha} \nu_n(id).$$

This is again a geometric series converging if $\alpha > 0$ implying that

$$\nu_\infty(id) = \mathrm{E}[Y] \left(1 + \frac{\lambda}{\alpha}\right) \qquad \text{if } \mathrm{E}[Y] < \infty.$$

We have now succeeded in finding $\kappa = \nu_\infty$ when the pieces are deterministic such that the renewal framework gives us a stationary process.

It is a bit surprising that there exists a stationary distribution even as $\alpha \to 0$, where $\varphi_x(s) = xe^{-\alpha s}$ resembles the step function more and more as α gets closer to 0.

3.2.3. Other deterministic pieces

Are there other possible choices of $\varphi_s(x)$? Besides the semigroup-requirement (1.62) such that the PDP is Markovian, the induction start works only if $\varphi_s(x) > 0$ all the

way up to the first jump. The jumps being from a Poisson process can happen at an arbitrary point of time with positive probability which implies

$$\varphi_s(x) > 0 \qquad \text{for all } s \text{ and } t.$$

The induction step merely relies on—other than $Y \geq 0$ a.s.—the fact that the functions are monotone in x such that for $x \leq x'$

$$\varphi_s(x) \leq \varphi_s(x') \qquad \text{for all } s.$$

Finally, in order to stay away from infinity

$$s \mapsto \varphi_s(x) \qquad \text{must be decreasing for } s \in \mathbb{R}$$

$\varphi_s(x) = xe^{bs} - \frac{a}{b}(1 - e^{bs})$ with $a > 0$ and $b < 0$ is the mean of CIR/OU but is not a PDMP piece that is it does not fulfill (1.62).

3.3. Diffusion pieces

So far the setup considered has assumed $Y \geq 0$ a.s. and the $\mathbf{X}^{(i)}$'s to be deterministic and strictly positive. We maintain positive Y's, but try to generalize the above by considering diffusion pieces not necessarily positive and want to investigate when a stationary distribution for the jump-diffusion exists.

We first establish that ν_∞ exists and the argument splits into two:

1. Show that $\nu_1 \preceq \nu_2 \preceq \ldots$, and
2. that ν_∞ has no mass at ∞, that is ν_∞ is a measure on $]0, \infty[$.

Theorem 3.3 gives precise condition for when 1. holds and relies on Lemma 3.2. Theorem 3.4 offers sufficient conditions for 2.

Lemma 3.2 (Domination Lemma)**.** *Let* $\mathbf{X}$ *and* $\widetilde{\mathbf{X}}$ *be diffusions on* $]l, r[$ *such that*

$$\begin{aligned} dX_t &= b(X_t)dt + \sigma(X_t)dB_t, & X_0 &= x_0, \\ d\tilde{X}_t &= b(\tilde{X}_t)dt + \sigma(\tilde{X}_t)dB_t, & \tilde{X}_0 &= \tilde{x}_0, \end{aligned}$$

and both $\mathrm{E}_{x_0}[X_t]$ *and* $\mathrm{E}_{x_0}[\tilde{X}_t]$ *exist for all* $t \geq 0$*. If* $x_0 < \tilde{x}_0$ *then*

$$P_t f(x) = \mathrm{E}_{x_0}[f(X_t)] \leq \mathrm{E}_{\tilde{x}_0}[f(\tilde{X}_t)] = P_t f(\tilde{x}), \qquad t \geq 0 \tag{3.12}$$

for f *bounded and non-decreasing.*

Proof. By assumption $\tilde{X}_0 > X_0$ from which it follows that

$$\tau := \inf\{t \geq 0 : X_t = \tilde{X}_t\} > 0$$

with the convention $\inf\{\emptyset\} = \infty$. On the set $\tau = \infty$ (3.12) holds as f is non-decreasing. What remains is to consider the set $\tau < \infty$.

Assume therefore $\tau < \infty$ and consider the τ-shifted $\mathbf{X}$-process

$$\begin{aligned} X_{\tau+t} &= X_\tau + \int_\tau^{\tau+t} b(X_{\tau+s})\, ds + \int_\tau^{\tau+t} \sigma(X_{\tau+s})\, dB_s \\ &= X_\tau + \int_0^t b(X_s)\, ds + \int_0^t \sigma(X_s)\, dB_s^* \end{aligned} \tag{3.13}$$

where $B^*_s := B_{\tau+s} - B_\tau$. The Brownian motion exhibits the strong Markov property and hence $\mathbf{B}^* = (B^*_s)_{s\geq 0}$ is again a Brownian motion and the SDE (3.13) therefore describes a diffusion. Similarly, the τ-shifted $\widetilde{\mathbf{X}}$ satisfies

$$\tilde{X}_{\tau+t} = \tilde{X}_\tau + \int_0^t b(\tilde{X}_s)\,ds + \int_0^t \sigma(\tilde{X}_s)\,dB^*_s. \tag{3.14}$$

Now for the core of the argument: Per definition of τ it holds that $X_\tau = \tilde{X}_\tau$ and hence (3.13) and (3.14) describe the very same SDE and Theorem 1.7 implies that $(X_{\tau+t})_{t\geq 0}$ and $(\tilde{X}_{\tau+t})_{t\geq 0}$ are indistinguishable. In particular $X_{\tau+t}(\omega) = \tilde{X}_{\tau+t}(\omega)$ for every ω and $t \geq 0$ implying (3.12). □

Important remark The lemma makes no assumptions on the nature of $]l, r[$ and in particular allows for diffusion attaining negative values! Since $y \geq 0$ it furthermore implies that $x \mapsto h_f(x,y) = E_x[f(X^{(1)}_{T_1} + Y_1]$ is non-decreasing for any f non-decreasing regardless of the diffusion $\mathbf{X}^{(1)}$ as long as the expectation exists. We state and proof:

Theorem 3.3. *Let the $\mathbf{X}^{(i)}$'s be any (!) diffusion with a stationary distribution and where $\mathrm{E}[X_t]$ exists for all t. Let the jumps be governed by a Poisson process with intensity λ. Then ν_∞ exists and is a measure on $]-\infty, \infty]$.*

Proof. Throughout let f be a continuous and non-decreasing function while η is the stationary distribution of the diffusion $\mathbf{X}^{(i)}$.

Recall from (3.2) that

$$\nu_{n+1}(f) = \iint \mathrm{E}_x[f(X^{(1)}_{T_1} + y)]\,d\nu_n(x)\,dG(y), \qquad n \geq 1.$$

It is shown by induction that $\nu_1 \preceq \nu_2 \preceq \ldots$. Put $\nu_1 \equiv \eta$. Stationarity then implies

$$\int \mathrm{E}_x[\varphi(X^{(1)}_s)]\,d\nu_1(x) = \int \varphi(x)\,d\nu_1(x)$$

for any φ continuous and bounded, where it should be noted that the s dependency disappears. A Poisson process with intensity λ governs the jumps and hence for any continuous and bounded φ it follows that

$$\begin{aligned}
\nu_2(\varphi) &= \iiint_0^\infty \mathrm{E}_x[\varphi(X^{(1)}_s + y)]\lambda e^{-\lambda s}\,ds\,d\nu_1(x)\,dG(y) \\
&= \iiint_0^\infty \varphi(x+y)\lambda e^{-\lambda s}\,ds\,d\nu_n(x)\,dG(y) \\
&= \iint \varphi(x+y)\,d\nu_1(x)\,dG(y)
\end{aligned}$$

Hence for f—since $Y \geq 0$ a.s. and therefore $f(x+y) \geq f(x)$ a.s.—it is seen that

$$\nu_2(f) \geq \iint f(x)\,d\nu_1(x)\,dG(y) = \nu_1(f)$$

equivalent to $\nu_1 \preceq \nu_2$ and proving the induction start.

The induction step follows the lines of the deterministic case closely. Assume $\nu_{n-1} \preceq \nu_n$. Lemma 3.2 gives that $x \mapsto h_f(x,y)$ is a non-decreasing function and hence

$$\nu_{n-1}(h_f(x,y)) \leq \nu_n(h_f(x,y))$$

equivalent to

$$\nu_n(f) \leq \nu_{n+1}(f)$$

as desired. Since the ν_n's are monotonically increasing ν_∞ exists as a probability on $]-\infty,\infty]$. □

Consider our favorite diffusion process the standard Brownian motion, which fulfills (3.12) of Lemma 3.2, but cannot be made stationary. It has mean 0 for any t and thus $\nu_1(id) = \mathrm{E}[Y]$. The second piece of the jump-diffusion is then a standard Brownian motion started at $\mathrm{E}[Y]$ and thus $\nu_2(id) = 2\,\mathrm{E}[Y]$. As $Y \geq 0$ a.s. this obviously implies that $\nu_\infty = \varepsilon_\infty$. This, so to speak, is a diffusion analog to the deterministic step process.

We therefore need to come up with some diffusion that, phrased informally, reverts down and away from ∞ with sufficiently high probability. Theorem 3.3 makes this statement precise: The diffusions considered must have a stationary distribution.

Knowing that ν_∞ exists we wish to investigate when it becomes a probability on $\mathbb{R}$ (excluding ∞!). One class of diffusions for which this holds is characterized in the following theorem.

Theorem 3.4. *Let the $\mathbf{X}^{(i)}$'s be any stationary diffusion fulfilling the requirement of Theorem 3.3. Let the jumps be Poisson with constant intensity $\lambda > 0$. Suppose $\alpha(t)$ and $\beta(t)$ exist such that*

$$\mathrm{E}_x[X_t^{(i)}] = \alpha(t)x + \beta(t), \quad |\alpha(t)| < 1 \quad \textit{for all } t, \quad \textit{and} \quad \int |\beta(t)|\,dt < \infty. \tag{3.15}$$

If in addition $\mathrm{E}[Y] < \infty$ then ν_∞ is a probability on $(-\infty,\infty)$ with expectation

$$\nu_\infty(id) = \frac{\mathrm{E}[Y] + \int_0^\infty \lambda e^{-\lambda s}\beta(s)\,ds}{1 - \int_0^\infty \lambda e^{-\lambda s}\alpha(s)\,ds}. \tag{3.16}$$

Proof. Theorem 3.3 gives that ν_∞ and $\mathrm{E}_x[X_t]$ exists such that (3.15) is meaningful. With $\mathrm{E}_x[X_t]$ having the required form of (3.15) one gets

$$\begin{aligned}\mathrm{E}_x[X_{T_1}^{(1)}] &= \int_0^\infty \lambda e^{-\lambda s}\,\mathrm{E}_x[X_s^{(1)}]\,ds \\ &= x\underbrace{\int_0^\infty \lambda e^{-\lambda s}\alpha(s)\,ds}_{=:\bar\alpha} + \underbrace{\int_0^\infty \lambda e^{-\lambda s}\beta(s)\,ds}_{=:\bar\beta}.\end{aligned}$$

Thus $\mathrm{E}_x[X_{T_1}^{(1)}] = \bar\alpha x + \bar\beta$ with $|\bar\alpha| < 1$ and $|\bar\beta| < \infty$ and therefore

$$\begin{aligned}\nu_{n+1}(id) &= \int_0^\infty\int E_x(X_{T_1}^{(1)} + y)\,d\nu_n(x)\,dG(y) = \mathrm{E}[Y] + \int E_x[X_{T_1}^{(1)}]\,d\nu_n(x) \\ &= \mathrm{E}[Y] + \bar\beta + \bar\alpha\,\nu_n(id).\end{aligned}$$

such that $\big(\nu_n(id)\big)_0^\infty$ is a finite geometric series with (3.16) as limit. □

Remark Theorem 3.4 can be expanded slightly. (3.15) states $\int |\beta(t)|dt < \infty$ but can be relaxed to requiring that $\int \beta(t)dt$ exists thus allowing for

$$\int \beta(t)dt = \begin{cases} -\infty \\ \infty \end{cases} .$$

The quantity of interest is $\bar{\beta} = \int_0^\infty \lambda e^{-\lambda t}\beta(t)\,dt$ which can now attain the values in $[-\infty, \infty]$. $|\bar{\beta}| < \infty$ was used in the proof and if this continuous to be the case the conclusion of the theorem sticks: ν_∞ is a probability measure on $\mathbb{R}$ with expectation as in (3.16). If in contrast $|\bar{\beta}| = \infty$ (3.16) is still valid with $\nu_\infty(id) = \bar{\beta}$ and thus ν_∞ is a probability on either $[-\infty, \infty)$ or $(-\infty, \infty]$ (concisely: on $\mathbb{R} \cup \{\bar{\beta}\}$).

Albeit seemingly artificial assuming whatever is needed to make the technique work, the x-affine structure of the expectation is a useful criterion. The standard Brownian motion with $\alpha(t) \equiv 1$ and $\beta(t) \equiv 0$ barely fails the requirements of the theorem, but it holds for the Cox-Ingersoll-Ross and the Ornstein-Uhlenbeck process as their expectation is given by (see (1.46))

$$\mathrm{E}_x[X_t] = xe^{bt} - \frac{a}{b}(1 - e^{bt}).$$

The geometric Brownian motion has expectation (see (1.40))

$$\mathrm{E}_x[X_t] = xe^{rt}$$

and thus also satisfies the theorem. Remember though, that there are certain constraints on the parameters in for the CIR and GBM to have a stationary distribution, see Section 1.3.4.

3.4. Negative jumps

Throughout we have assumed the jump sizes were non-negative, that is $Y \geq 0$ a.s., but as the following simple observation shows this is not a very limiting assumption when considering applications. Section 4.2 offers an example of a model with negative jumps.

Observation 3.5. *If* $\mathbf{X}$ *with values in* $]l, r[$ *is a stationary process and* $S :]l, r[\to \mathbb{R}$ *is a strictly monotone map then* $S(\mathbf{X})$ *is stationary as well.*

Naturally, this observation applies to any kind of stationary process be it a diffusion or jump size/target jump-diffusion. It is particularly useful in the jump size setting as the model of stationary jump-diffusions with negative jumps are easily produced.

In order to get a model with negative jumps all that needs to be done is producing a stationary $\mathbf{X}$ and then use a S that is monotonically decreasing. Then any positive jump for $\mathbf{X}$ results in a negative jump for $S(\mathbf{X})$.

3.5. Discussion of results

Albeit having some extent of success in establishing quite a few results about when a stationary distribution exists in the jump size models, it has not been possible to identify an explicit stationary density.

In Chapter 2 we were able to derive an expression for the density in some cases when the intensity was constant, while we for non-constant intensities could provide a criterion a stationary density had to fulfill.

In the setting of this chapter we still know that a stationary density exists when the jumps stem from a Poisson process, but so far we have not been able to find its density. This is even though the analysis of the jump size model has been linked to a jump target model with target distribution ν_∞.

The key in Chapter 2 to do so was knowing the target distribution. When instead specifying the jump sizes its equivalent becomes ν_∞^λ which—unlike in the jump target model—depends on the jump-intensity $\lambda(x)$! Disregarding the fact that finding ν_∞^λ has not even in the most simple of settings been possible, this alone complicates matters enormously. For instance this makes an explicit form like (2.7) unattainable.

In the jump target model the jump target distribution was given beforehand and independent of all other quantities involved; as nice as it gets. On the other hand in the jump size setting ν_∞ depends complete on all three entities of the model: the diffusion pieces $\mathbf{X}^{(i)}$, the jump-intensity $\lambda(x)$, and the jump size distribution.

and since, as we shall see, it is very difficult to determine ν_∞. This means that while an explicit expression for the stationary density we are not able to derive it. So far all we have been able to say about ν_∞ is that it is a genuine probability on $(-\infty, \infty)$.

The distribution of ν_∞ is the heart of the jump size model and also absolutely essential when we for x-dependent intensity $\lambda(x)$ we wish to establish that the existence of a stationary distribution, that is the first jump exists with positive probability. In order for the jump size model to become truly useful more time and energy would have to be spend on methods of determining ν_∞.

While it has not been possible to derive an expression for v_∞ the density of ν_∞ the chapter nevertheless contains some tools that can be used to numerically find the density and possible even a closed expression.

We present the idea in a simplified context. Consider the deterministic case $\varphi_s(x) = xe^{-\alpha s}$ with $\alpha > 0$, or equivalently the GBM, with degenerate jumps of size a. From Section 3.2 we know that if $\mathbf{X}^{(1)}$ is started with density κ then the density η for X_{T_1} is given by (3.8). Hence the density of $X_{T_1}^{(1)}$ would be

$$\eta(z) = \frac{\lambda}{\alpha} z^{\lambda/\alpha - 1} \int_z^\infty w^{-\lambda/\alpha} \kappa(w)\, dw,$$

with the density of $X_{T_1}^{(1)} + Y$ being

$$v_1 = \eta(z - a).$$

This can be used to recursively determine v_∞ by putting $\kappa = v_1$ to obtain an expression for v_2 and then iterate.

The link between v_{n+1} and v_n for $n \geq 1$ in the general case when Y has density g becomes

$$v_{n+1} \equiv \left(g * \frac{\lambda}{\alpha} y^{\lambda/\alpha - 1} \int_y^\infty w^{-\lambda/\alpha} v_n(w)\, dw \right). \tag{3.17}$$

In Chapter 2 the infinitesimal generator was used to find a integro-differential equation for the stationary density with great success. Considering the above difficulties and the devilish interconnection of ν_∞ with $\mathbf{X}^{(i)}$, $\lambda(x)$, and Y, it would not surprise us if the analysis breaks down.

This is in fact the case and with the generator (2.5) in this model being

$$\begin{aligned} Af(x) &= \frac{\sigma^2(x)}{2} f''(x) + b(x) f'(x) + \lambda(x) \int f(x+y) - f(x) v_\infty(y)\, dy \\ &= A^d f(x) + \lambda(x) \int f(x+y) v_\infty(y)\, dy - \lambda(x) f(x). \end{aligned}$$

as the process jumps from x to $x+y$. Even when keeping jump intensities constant $\lambda(x) \equiv \lambda$, we head into trouble and with u being the density of ν find

$$\begin{aligned} \nu(Af) &= \int A^d f u\, dx + \lambda \int \left(\int f(x+y) v_\infty(y)\, dy \right) u(x) - f(x) u(x)\, dx \\ &= \int A^d f u - \lambda f u\, dx + \lambda \iint f(x+y) v_\infty(y) u(x)\, dy\, dx \end{aligned}$$

for $f \in C_0^2(\mathbb{R})$. Even when pretending v_∞ to be known this seems unlikely to be fruitful as it is simple not possible to separate f in the double integral in any successful way.

CHAPTER 4

Applications

This chapter presents applications of the theory developed in Chapters 2 and 3. As is clear from the discussion in Section 1.3.4 it is very difficult to apply the jump size model. Nevertheless a model is formulated to give the reader a taste of what phenomena might be modeled using jump sizes.

Both models exhibit non-constant jump intensity and even in the case of the target model this leads to equations not solvable without the numerical analysis. The application of the target model is described in detail and the numerical analysis is initiated; a full analysis is too complex to fit within the span of this thesis.

The models were suggested by Susanne Ditlevsen.

4.1. Models using jump targets

We build a neuron model in which the membrane of the neuron holds a certain electrical potential and can "fire," that is release, this potential as part of neuronal activity. Such a model is called an *integrate-and-fire* neuronal model.

The potential is modeled as a continuous-time stochastic process here a jump-diffusion. A threshold $a > 0$ is defined such that while the potential being less than a the probability for a release (jump) is tiny but as soon as the potential passes a the probability increases drastically approaching 1. An intensity modeling this kind of behavior is

$$\lambda(x) = \exp\left(\frac{x-a}{b}\right), \qquad 0 < b \ll a. \tag{4.1}$$

The smaller b is the more drastic is the change of the jump intensity around a from tiny to huge. In the limit $b \to 0$ we would expect a pure threshold behavior.

The diffusion parts of the jump-diffusion is described by an OU-process , that is the SDE

$$dX_t^{(i)} = -\beta(X_t^{(i)} - \alpha)dt + \sigma dB_t, \qquad \beta, \sigma > 0,\ \alpha \in \mathbb{R}.$$

The jump target distributions Y_i's and the boundary condition for the above SDE must model the following behavior. When the neuron fires there might be a small residual potential left or the potential might become slightly negative due to the

surrounding neurons. To model this we let the Y_i's follow a normal distribution with mean 0 and variance $0 < \tau \ll a$, that is

$$Y_i \overset{\mathcal{D}}{=} \mathcal{N}\left(0, \tau^2\right)$$

To interpret the parameters it pays to recall (1.43) describing the stationary distribution of the OU-process. It is a $\mathcal{N}\left(\alpha, \frac{\sigma^2}{2\beta}\right)$ making α easy to interpret since it simply is the level around which the OU-process hovers in the long term. β and σ together determine the variance of the stationary distribution and a larger β ensures that the process reaches the level of α more rapidly.

Having the model in place the first question to ask is whether a stationary distribution exists or not. According to Theorem 2.1 we must check if

$$\mathrm{E}_G[T_1] < \infty.$$

The intensity considered here is non-constant, which complicates matters significantly. Fortunately Corollary 2.2 is at our disposal and leaves a way out. We replace $\lambda(x)$ from (4.1) with

$$\widetilde{\lambda}(x) := \min\left(\exp\left(\frac{x-a}{b}\right), \varepsilon\right), \qquad 0 < b \ll a \tag{4.2}$$

for some very small $\varepsilon > 0$. As $\widetilde{\lambda}(x)$ is bounded below Corollary 2.2 immediately gives that a stationary distribution exists.

This of courses changes the model we just so proudly build and what consequences does this mangling with the intensity have? From a modeling point of view it is rather insignificant as ε can be chosen however small, but admittedly $\widetilde{\lambda}$ is algebraically less attractive. This however is of little concern when first we set out to find the stationary distribution as the only feasible way to do so is utilizing criterion (2.25). To compare the two we first discuss (2.25) for $\lambda(x)$.

In case of $\lambda(x)$, (2.25) becomes

$$\begin{aligned} \frac{\sigma^2}{2}u''(x) + \beta(x-\alpha)u'(x) + \left[\beta - \exp\left(\frac{x-a}{b}\right)\right]u(x) \\ = -\frac{1}{\sqrt{2\pi\tau^2}}\exp\left(\frac{-x^2}{2\tau^2}\right) \cdot \int \exp\left(\frac{y-a}{b}\right) u(y)\, dy \end{aligned} \tag{4.3}$$

with boundary condition u a probability density. Obviously, this is by no means a trivial problem to solve since at least two circumstances complicate matters. First of all the functional equation involves an integral over the entire axis rather than from say 0 to x in which case there might be hope to solve (4.3) using Laplace transforms, see [Fe66] p. 441-445. Second of all the boundary condition is highly non-standard as we do not know any specific values, but are simply told that u has to be a density. This troubles us when attempting to find a solution both by algebraic and numerical means.

Regardless of whether one takes the analytical or numerical approach it is fruitful to formulate (4.3) as a two-step problem. Noting that the integral although depending on u is a constant say $K > 0$, we can begin by considering

$$\frac{\sigma^2}{2}v_K''(x) + \beta(x-\alpha)v_K'(x) + \left[\beta - \exp\left(\frac{x-a}{b}\right)\right]v_K(x) = K \cdot \frac{1}{\sqrt{2\pi\tau^2}}\exp\left(\frac{-x^2}{2\tau^2}\right),$$

which is a regular second order inhomogeneous ODE. Suppose we can somehow find a family of solutions $(v_K)_{K\in\mathbb{R}_+}$ it must include the desired density u. A necessary condition for this v_K is to satisfy

$$\int \exp\left(\frac{y-a}{b}\right) v_K(y)\,dy = K$$

together with $\int v_K\,dy = 1$ and $v_K \geq 0$.

Even this second order inhomogeneous ODE is not solvable algebraically[1] and we are thus forced into a numerical procedure. For such it is irrelevant if $\lambda(x)$ or $\widetilde{\lambda}(x)$ is used and hence the ODE to solve numerically becomes

$$\frac{\sigma^2}{2}v_K''(x) + \beta(x-\alpha)v_K'(x) + \left[\beta - \widetilde{\lambda}(x)\right] v_K(x) = -K \cdot \frac{1}{\sqrt{2\pi\tau^2}} \exp\left(\frac{-x^2}{2\tau^2}\right) \tag{4.4}$$

with the function we are looking for fulfilling

$$\int \widetilde{\lambda}(y) v_K(y)\,dy = K \tag{4.5}$$

and $\int v_K\,dy = 1$ and $v_K \geq 0$.

It has therefore been argued that both from a modeling perspective and when solving the problem it is natural to use $\tilde{\lambda}(x)$ rather than $\lambda(x)$.

In order to find v_K we, a priori, need two things: to iterate and a bit of luck! The numerical procedures familiar to the author all require a starting point c at which $v_K(c)$ and $v_K'(c)$ is specified. In this setting $v_K(c) = v_K'(c) = 0$ seems like a safe bet for c sufficiently negative depending on the choice of parameters ($c = -1000\tau$ should do). Fixing K at some arbitrary positive value such as 1 we can find a first candidate: v_1. This then needs to be normalized (hoping for $v_1(x) \geq 0$) returning $\widetilde{v}_1$ and (4.5) now yields a new constant K'. Thus we run the procedure again with K' in place of K and keep our fingers crossed that the density produced at some point converges.

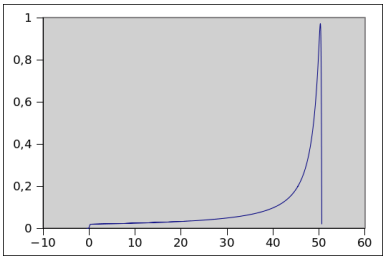

Figure 4.1.: Plot of ψ, a numerical solution to (4.4) with parameters $a = 50$, $\alpha = 50$, $b = 0.1$, $\beta = 1$, $K = 1$, $\sigma = 2$, and $\tau = 0.1$. ψ is non-negative and integrates to approximately 4.67.

Time being short the author was not able to write such a procedure, but attempts were made producing a function ψ plotted in Figure 4.1. No iterations were done and therefore ψ has not been shown to converge and note that ψ does not integrate to 1. Despite this the plot is interesting since it fits with the intuition of the model.

The threshold $a = 50$ with the tiny $b = 0.1$ is clearly seen in the plot: as soon as the process comes above 50 a jump is triggered and the density rapidly falls to 0.

[1] At least not by methods familiar to the author at present.

$\alpha = 50$ ensures that jumps indeed occurs as the OU-process' expectation is $\alpha = 50$ and with $\beta = 1$ and $\sigma = 2$ it rather rapidly approaches $\alpha = 50$ after being reset to approximately 0. Following a jump the reset target is picked from a $\mathcal{N}(0, 0.1)$-distribution since $\tau = 0.1$ and as it is seen that very little mass is on the negative axis. In fact, at any point on the negative axis ψ is of order 10^{-3} or less.

Without rigorous experiments having been carried out, it should be remarked that the numerical analysis seems a bit instable. Nevertheless it is the gut feeling of the author from playing with the code and different parameter values that there is hope for finding the stationary density numerically. The initial plot of ψ being so intuitive supports this.

4.2. Models using jump sizes

As mentioned many times before the jump size models are close to hopeless at this state. Finding the stationary distribution or even showing its existence is very difficult and the non-constant jump intensity only adds to this frustration. This very brief section outlines a jump size model, which unfortunately cannot be analyzed to any extent and as such should rather be seen as inspiration and motivation for further research.

Recall Observation 3.5 stating that if $\mathbf{X}$ with values in $]l, r[$ is a stationary process and $S :]l, r[\to \mathbb{R}$ is a strictly monotone map then $S(\mathbf{X})$ is stationary as well. Section 3.4 already mentioned how this might be used to obtain a jump-diffusion with negative jumps and we formulate the following model.

The beta cell is found in the islets of Langerhans of the pancreas. It produces the hormone insulin continuously used for regulating the blood sugar levels. The release of insulin from a beta cell depends on the concentration of sugar in the blood and happens discretely in quanta. As such it is natural to model it using a jump-diffusion where the jump sizes are a decrease in insulin concentration with the beta cell.

Denote by Z_t the concentration of insulin in a beta cell at time t and as negative jumps are called for we use a monotonically decreasing transformation S to model $Z_t = S(X_t)$ where for example

$$dX_t = -\beta(X_t - \alpha)dt + \sigma X_t dB_t, \qquad \alpha\beta > 0,\ \frac{2\beta}{\sigma^2} > -1,$$

and either

$$S(X_t) := \frac{1}{X_t} \qquad \text{or} \qquad S(X_t) := \exp(-cX_t), \quad c > 0.$$

Note that if $\alpha = 0$ then X_t is a GBM. The jumps need to be positive and $Y \overset{\mathcal{D}}{=} \Gamma(\eta, \tau)$ is chosen as jump-distribution. An intensity function should be chosen such that the intensity is large when $\mathbf{Z}$ is large and small when $\mathbf{Z}$ is close to 0. A simple choice would be

$$\lambda(x) := \frac{1}{x}$$

$\lambda(x) = S(x)$, or any other function displaying the desired features.

APPENDIX A

Distribution theory, convolutions and integral transformations

This appendix contains a brief description of tools important in this thesis. The section on distribution theory is inspired by [Gr09]. Convolutions and integral transforms have analytical as well as probabilistic interest and a combined approach is taken here. Again, [Gr09] has provided inspiration, while [Fe66] has given fruitful insight into the probability theory. Fourier and Laplace transforms are ubiquitous in the literature and for instance described in [KT75].

A.1. Distribution theory

Some functions not differentiable in the ordinary sense can be assigned a meaningful derivative in a distributional sense, see [Gr09]. We list the following definition and results.

Definition A.1. A distribution on $\mathbb{R}$ is a continuous linear functional Λ on the space of test functions $C_0^\infty(\mathbb{R})$. The value of Λ for $\varphi \in C_0^\infty(\mathbb{R})$ is denoted by

$$\Lambda(\varphi) \text{ or } \langle \Lambda, \varphi \rangle \quad \text{with} \quad \langle \cdot, \cdot \rangle : L_{1,loc}(\mathbb{R}) \times C_0^\infty(\mathbb{R}) \to \mathbb{R},\ (f, \varphi) \mapsto \int_{\mathbb{R}} f\varphi dt. \quad \text{(A.1)}$$

The derivative of a distribution is given by

$$(\partial\Lambda)(\varphi) = \langle \partial\Lambda, \varphi \rangle = \int \partial\Lambda\varphi dt = [\Lambda\varphi]_{-\infty}^{\infty} - \int \Lambda\partial\varphi dt = -\langle \Lambda, \partial\varphi \rangle$$

and more generally

$$\langle \partial^n\Lambda, \varphi \rangle = (-1)^n \langle \Lambda, \partial^n\varphi \rangle . \quad \text{(A.2)}$$

Note, that ∂^n denotes derivatives in distributional sense, while $\frac{d^n}{dx^n}$ is used for regular differentiation.

Theorem A.1. *If a function f has a nth derivative on $\mathbb{R}_-$ and $\mathbb{R}_+$, respectively, its derivative in a distributional sense is*

$$v = \begin{cases} \frac{d^n}{dx^n} f & \text{on } \mathbb{R}_-, \\ \frac{d^n}{dx^n} f & \text{on } \mathbb{R}_+. \end{cases}$$

Proof. See Lemma 3.6 in [Gr09] p. 34. □

For example $\partial|x| = \operatorname{sgn} x$ (the sign function).

Theorem A.2. *The space $C_0^\infty(\mathbb{R})$ is dense in $L_p(\mathbb{R})$ for $p \in [1, \infty)$.*

Proof. See Theorem 2.15 part 3 in [Gr09] p. 22. □

A.2. Convolutions

Two types of convolutions are treated here. One, denoted by $*$ is an operation between two functions while the other denoted by $\star$ is an operation between a continuous function and a probability distribution function (more generally a bounded non-decreasing function).

Definition A.1. Let $f, g \in C(\mathbb{R}) : \mathbb{R} \to \mathbb{R}$ be subject to certain integrability constraints and define

$$(f * g)(x) = \int_{\mathbb{R}} f(y)g(x-y)dy. \tag{A.3}$$

Let F be a distribution function and g continuous, again such that the following integral is well-defined. Define

$$(g \star F)(x) = \int_{\mathbb{R}} g(x-y)dF(y). \tag{A.4}$$

If F has density f then $(g \star F) \equiv (g * f)$. Notation: f^{n*} and $F^{n\star}$ denotes the n-fold convolution.

Despite their similarities, it is very useful to have developed the notion of both convolutions when dealing with probability distributions; some with densities and some without. A myriad of choices for conditions on the involved functions/distributions such that the integrals are well-defined exists. In applications this is rarely a problem and one simply has to consider the specific uses.

Where $*$ is commutative and associative for all function for which it is defined, the $\star$-operation is not. Unless g itself is a distribution function $(F \star g)$ is meaningless. On the other hand, if G and F both are distribution functions the $\star$-operation is commutative and associative. In this case the $\star$-operation has a probabilistic interpretation as

$$P(X + Y \leq x) = (G \star F)(x)$$

where F and G are the distribution functions for X and Y. This interpretation of course translates to the $*$-operation provided densities exist. We list the following well-known properties:

- If G is a probability distribution function so is $F \star G$.
- If f and g are both probability densities so is $f * g$.
- Let F/f and G/g represent two probability distributions with atoms $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$. Then the atoms for the convolved distributions are exactly all possible sums $a_i + b_j$, $i, j \geq 1$. In particular if the distributions are continuous so is the convolved.

From Fubini's Theorem it follows that

$$\int_{\mathbb{R}} (f * g)(t)dt = \int_{\mathbb{R}} f dt \int_{\mathbb{R}} g dt. \tag{A.5}$$

A particular special case of interest arises when f and g both have support on $\mathbb{R}_-$ or $\mathbb{R}_+$:

$$(f * g)(t) = \int_0^t f(\tau)g(t-\tau)d\tau \qquad \text{if } f, g \in \text{supp}\left(\mathbb{R}_+\right)$$

and

$$(f * g)(t) = \int_{-t}^0 f(\tau)g(t-\tau)d\tau \qquad \text{if } f, g \in \text{supp}\left(\mathbb{R}_-\right).$$

The convolved function behaves more nicely than the two original functions and one habitually refers to convolution as a "smoothing operator." For instance, the convolved function is differentiable at a given point as many times as the two original functions are together.

A.3. Fourier transform

A popular choice of integral transform is the Fourier transformation closely related to the characteristic function of a real random variable. Denoted by either $\hat{f}$ or $\mathcal{F}(f)$, many popular choices for the definition of the (one-dimensional) Fourier transform exist. This thesis uses

$$\mathcal{F}\left(f\right)(s) = \hat{f}(s) = \frac{1}{\sqrt{2\pi}} \int e^{-ist} f(t)dt \tag{A.6}$$

in which case the inverse becomes [1]

$$f(t) = \frac{1}{\sqrt{2\pi}} \int e^{ist} \mathcal{F}\left(f\right)(s)ds. \tag{A.7}$$

We shall need the following facts:

$$\mathcal{F}\left(af(t) + bg(t)\right)(s) = a\hat{f}(s) + b\hat{g}(s) \tag{A.8}$$

$$\mathcal{F}\left(\partial^n f\right)(s) = (is)^n \hat{f} \tag{A.9}$$

$$\mathcal{F}\left(f * g\right)(s) = \sqrt{2\pi}\hat{f}(s)\hat{g}(s) \tag{A.10}$$

and a brief table of Fourier transforms

$$\mathcal{F}\left(e^{-a|t|}\right)(s) = \sqrt{\frac{2}{\pi}} \cdot \frac{a}{s^2 + a^2}, \qquad a > 0, \tag{A.11}$$

$$\mathcal{F}\left(e^{-a|t|-bt}\right)(s) = \sqrt{\frac{2}{\pi}} \cdot \frac{a}{s^2 + a^2 - b^2 - 2bis}, \qquad a > b > 0 \tag{A.12}$$

$$\mathcal{F}\left(1\right)(s) = \sqrt{2\pi}\delta_0. \tag{A.13}$$

All of the above except (A.12) are standard textbook references. For (A.12) we have

$$\sqrt{2\pi} \cdot \mathcal{F}\left(e^{-a|t|-bt}\right)(s) = \int_{-\infty}^0 e^{(a-b)t} e^{-ist} dt + \int_0^\infty e^{-(a+b)t} e^{-ist} dt$$

[1] The inverse is seen to be proportional to the characteristic function.

where both integrals converges simultaneously if and only if $a > b > 0$. When they do

$$\sqrt{2\pi} \cdot \mathcal{F}\left(e^{-a|t|-bt}\right)(s) = \frac{1}{a-b-is} + \frac{1}{a+b+is} = \frac{2a}{[a-(b+is)][a+(b+is)]}$$

yielding (A.12).

References

[Ja06] Jacobsen, Martin, *Point Process Theory and Applications, Marked Point and Piecewise Deterministic Processes*, Probability and its Applications, Birkhäuser, Boston, 2006.

[Ok00] Øksendal, B., *Stochastic Differential Equations*, Springer-Verlag Heidelberg New York, 5th edition, corrected printing, 2000.

[Ja08] Jacobsen, Martin, *Stokastiske integraler*, Afdeling for Anvendt Matematik og Statistik, University of Copenhagen, 1st edition, 2008.

[KT75] Karlin, S., Taylor, H.M., *A first course in stachastic processes*, New York Academic Press, 2nd edition, 1975.

[Fe66] Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. II, Wiley, New York, 1966.

[Ha09] Hansen, Ernst, *Measure theory*, Department of mathematical sciences, University of Copenhagen, 4th edition, 2009.

[Ja03] Jacobsen, Martin, *Videregående Sandsynlighedsregning*, Afdeling for Anvendt Matematik og Statistik, University of Copenhagen, 3rd edition, 2003.

[RW00] Rogers, L.C.G., Williams, D., *Diffusions, Markov processes and martingales*, Vol. 1-2, Cambridge University Press, reissued 2nd edition, 2000.

[IM65] Itô, K., McKean, H.P., *Diffusion processes and their sample paths*, Springer-Verlag, Berlin, 1965.

[Ro83] Roos, S.M., *Stochastic processes*, New York: John Wiley & Sons, 2nd edition, 1996.

[Mi72] Miller, D.R., *Existence of limits in regenerative processes*, The Annals of Mathematical Statistics, Vol. 43, No. 4, p. 1275-1282.

[BS02] Borodin, A.N., Salminen, P., *Handbook of Brownian motion: facts and formulae*, Probability and its Applications, Birkhäuser, Basel, 2nd edition, 2002.

[Gr09] Grubb, G., *Distributions and operators*, Springer, 1st edition, 2009.

Index