Åbo Akademi University

Bao Quoc Ta

Excessive Functions, Appell Polynomials and Optimal Stopping

Excessive Functions, Appell Polynomials and Optimal Stopping

Excessive Functions, Appell Polynomials and Optimal Stopping

Bao Quoc Ta

PhD Thesis in Applied Mathematics
Department of Natural Sciences
Åbo Akademi University
Åbo, Finland, 2014

ISBN 978-952-12-3045-5
Painosalama Oy – Turku, Finland 2014

Acknowledgement

First and foremost, I would like to express my deepest gratitude to my supervisor, Professor Paavo Salminen, not only for his guidance but also for his great patience, kind encouragement and constant support throughout my studies. It has been a privilege for me to be his PhD student.

My sincere thanks go to the pre-examiners, Professor Luis Alvarez from Turku School of Economics and Professor Boualem Djehiche from KTH Royal Institute of Technology for taking their time and efforts to review my thesis. I am also grateful to Professor Boualem Djehiche for accepting to be the opponent in the public defence of my thesis.

I want to thank the board of the Finnish Doctoral Program in Stochastics and Statistics for organizing a lot of interesting scientific activities and creating a warm stimulating atmosphere for doing research, especially during the annual summer schools. I am also thankful to Doctor Sören Christensen from University of Kiel for his collaboration and many interesting discussions. I wish to thank to my colleagues at the Department of Mathematics, Åbo Akademi University for a pleasant and friendly environment.

Financial support from the Academy of Finland project 127719, the Finnish Doctoral Program in Stochastics and Statistics and the Rektors stipendium of Åbo Akademi University is gratefully acknowledged.

Finally, I would like to thank my family for their love and for always being beside me.

Åbo, April 2014

Bao Quoc Ta

On the thesis

This thesis consists of an introduction and four original research articles:

[I] S. Christensen, P. Salminen and B. Q. Ta: *Optimal stopping of strong Markov processes,* Stochastic Process. Appl. 123(2013), no 3, 1138-1159.

[II] P. Salminen and B. Q. Ta: *Differentiability of excessive functions of one-dimensional diffusions and the principle of smooth fit*, preprint available in arXiv:1310.1901.

[III] B. Q. Ta: *Probabilistic Approach to Appell polynomials*, preprint available in arXiv:1307.4431

[IV] B. Q. Ta: *A note on the generalized Bernoulli and Euler polynomials*, Eur. J. Pure Appl. Math, 6(2013), no 4, 405-412.

Author's contributions to Articles [I] and [II]

[I] All authors contributed to the planning and writing the article. The main result is due to PS. The proof is joint work of PS and BT. The extension to the two-sided case and the further developments are due to SC. Examples are joint work of the authors.

[II] The article is joint work of the authors. The main idea is from PS.

Abstract

The main topic of the thesis is optimal stopping. This is treated in two research articles. In the first article we introduce a new approach to optimal stopping of general strong Markov processes. The approach is based on the representation of excessive functions as expected suprema. We present a variety of examples, in particular, the Novikov-Shiryaev problem for Lévy processes. In the second article on optimal stopping we focus on differentiability of excessive functions of diffusions and apply these results to study the validity of the principle of smooth fit. As an example we discuss optimal stopping of sticky Brownian motion. The third research article offers a survey like discussion on Appell polynomials. The crucial role of Appell polynomials in optimal stopping of Lévy processes was noticed by Novikov and Shiryaev. They described the optimal rule in a large class of problems via these polynomials. We exploit the probabilistic approach to Appell polynomials and show that many classical results are obtained with ease in this framework. In the fourth article we derive a new relationship between the generalized Bernoulli polynomials and the generalized Euler polynomials.

Sammanfattning

Huvudtemat i avhandlingen är optimal stopping. Detta behandlas i de två första artiklarna. I den första artikeln presenteras en ny approach till optimal stopping av starka Markov processer. Denna baserar sig på en representation av excessiva funktioner som väntevärdet av maximet av en funktions värde evaluerat vid processens tillstånd. För att påvisa metodens användbarhet behandlas ett antal exempel, speciellt diskuteras Novikov-Shiryayev problemet för Lévy processer. I den andra artikeln studeras differentierbarheten hos excessiva funktioner av en-dimensionella diffusionsprocesser. De erhållna resultaten används för att undersöka smooth-fit principen i optimal stopping. Som ett exempel analyseras optimal stopping av en klibbig (sticky) brownsk rörelse. Den tredje uppsatsen i avhandlingen är en översikt av Appell polynom genererade medelst stokastiska variabler. Novikov och Shiryayev upptäckte att Appell polynom har en viktig roll i optimal stopping av Lévy processer och kunde karakterisera optimala stopping regler för en stor klass av problem för Lévy processer med hjälp av Appell polynom. I uppsatsen om Appell polynom används sannolikhetsteoretiska metoder för

att bevisa många klassiska egenskaper av dessa polynom men det härleds också nya resultat t.ex. om väntevärdesrepresentationer. Speciellt behandlas Bernoulli, Euler, Hermite och Laguerre polynom. I den fjärde artikeln presenteras ett nytt samband mellan generaliserade Bernoulli och Euler polynom.

Contents

1 Foreword

In general, the objective in an optimal stopping problem (OSP) is to find a stopping time at which the underlying stochastic process should be stopped so that the expectation of a given reward function reaches the maximum at the stopping time. There are many more or less practical situations where optimal stopping problems appear. The so-called secretary problem is a classical and much studied optimal stopping problem in discrete time. However, the origin of the theory of optimal stopping is in mathematical statistics, more specifically, in sequential analysis initiated by A. Wald and J. Wolfowitz [57] in late forties and at the beginning of fifties. A new application of the theory of optimal stopping was found by Merton [37] in seventies who showed that the problem of finding a fair price of an American option is, in fact, an optimal stopping problem. We remark also that the optimal stopping and singular control problem are closely related, see, e.g., [2, 30, 31].

In the first research article of the thesis we are concerned with optimal stopping problems for general strong Markov processes. We present a new approach which can be used to solve the problems for Hunt processes, in particular, for Lévy processes and diffusions. The main ingredient of the approach is to use the representation of an α-excessive function as expected suprema.

In optimal stopping problems for one-dimensional diffusions, the so called principle of smooth fit is a much used tool to find solutions. However, the most studies focus only on the diffusions with differentiable scale functions and speed measures absolutely continuous with respect to the Lebesgue measure. In the second article we are, in particular, interested in studying differentiability of excessive functions of linear diffusions with general scale functions and speed measures. These results are then applied to study the validity of smooth fit in optimal stopping problems of linear diffusions.

As noticed by Novikov and Shiryaev [39] the Appell polynomials play an important role in solving optimal stopping problems. The crucial property hereby is that an Appell polynomial associated with running maximum of a Lévy process stopped at an exponential time has a unique positive root. Besides optimal stopping problems, Appell polynomials have a wide range of applications in, e.g., constructing time-space martingales for Lévy processes [50, 54] and finding probability distribution of the ruin time in insurance mathematics (see, e.g., [43, 44]).

In the third and fourth research articles we utilize the probabilistic approach to Appell polynomials and derive many of their properties. It is seen that the probabilistic approach provides us with a deeper understanding of the behaviour of the Appell polynomials. We present new proofs of classical results but also derive new results, e.g., the moment representations of Appell polynomials and relationships between the generalized Bernoulli polynomials and the generalized Euler polynomials.

2 Markov processes

2.1 Basic definitions

Let $X = (X_t)_{t\geq 0}$ be a stochastic process on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ taking values in the measurable space $(E, \mathcal{E})$. It is assumed that X_t is $\mathcal{F}_t$-measurable. Denote by $\mathbb{P}_x$ and $\mathbb{E}_x$ the probability and expectation of the process starting at $x \in E$, respectively. We refer to [9, 11, 13, 17, 46] for the general theory of Markov processes. Here, we will restrict our attention to time-homogeneous Markov processes.

Definition 2.1. *For each $t \geq 0$, the function $P_t : E \times \mathcal{E} \to [0,1]$ is a time-homogeneous transition function if*

(i) for each $A \in \mathcal{E}$ and $t \geq 0$ the map $x \to P_t(x, A)$ is measurable,

(ii) for each $x \in E$ and $t \geq 0$ the map $A \to P_t(x, A)$ is a measure,

(iii) for each $x \in E$ and $A \in \mathcal{E}$,

$$P_0(x, A) = \delta_{\{x\}}(A) \quad (\textit{unit mass at } x),$$

(iv) for each $x \in E, A \in \mathcal{E}$ and $t, s \geq 0$, the Chapman-Kolmogorov equation holds

$$P_{t+s}(x, A) = \int_E P_t(x, dy) P_s(y, A).$$

Definition 2.2. *The process X is called a time-homogeneous Markov process with respect to the filtration $\{\mathcal{F}_t\}$ and having the transition function P_t if*

$$\mathbb{P}_x(X_{t+s} \in A | \mathcal{F}_s) = P_t(X_s, A), \qquad \mathbb{P}_x - a.s., \tag{1}$$

for all $s, t \geq 0$, $A \in \mathcal{E}, x \in E$.

Equality (1) is also often written as

$$\mathbb{P}_x(X_{t+s} \in A|\mathcal{F}_s) = \mathbb{P}_{X_s}(X_t \in A), \quad \mathbb{P}_x - a.s.$$

Intuitively speaking, for a Markov process X the prediction of the behaviour of X in the future depends only on its current state and does not depend on the past history of the process.

Let f be a bounded and measurable function on E and define

$$P_t f(x) := \int_E f(y) P_t(x, dy).$$

It is seen that $P_t f$ is also bounded and measurable on E. Moreover, the operator P_t is a semigroup, i.e.,

$$P_{t+s} = P_t P_s.$$

It follows from (1) that for every bounded, measurable function f it holds

$$P_s f(X_t) = \mathbb{E}_x(f(X_{t+s})|\mathcal{F}_t), \quad \mathbb{P}_x - a.s.$$

Definition 2.3. *Let $\lambda > 0$ and f a bounded and measurable function on E. For each $x \in E$, define*

$$U^\lambda f(x) := \mathbb{E}_x\Big(\int_0^\infty e^{-\lambda t} f(X_t) dt \Big) = \int_0^\infty e^{-\lambda t} P_t f(x) dt.$$

The operator U^λ is called the λ-potential or λ-resolvent of X.

Definition 2.4. *A random variable τ taking values on $\mathbb{R}^+ \cup \{+\infty\}$ is called a stopping time with respect to $\{\mathcal{F}_t\}$ if $\{\tau \le t\} \in \mathcal{F}_t$ for all $t \ge 0$.*

In case the filtration $\{\mathcal{F}_t\}$ is right continuous, i.e., $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$, then τ is a stopping time if and only if $\{\tau < t\} \in \mathcal{F}_t$ for all t. Some important examples of stopping times are the first hitting time $H_B := \inf\{t > 0 : X_t \in B\}$ and the first entrance time $\tau_B := \inf\{t \ge 0 : X_t \in B\}$, where under very general conditions, B can be taken to be an arbitrary Borel subset of E. It is seen that when $x \notin B$ then the times H_B and τ_B are the same.

An important σ-algebra associated with a stopping time τ is defined via

$$\mathcal{F}_\tau := \{A \in \mathcal{F} : A \cap \{\tau \le t\} \in \mathcal{F}_t, \quad \text{for all} \quad t \ge 0\}.$$

This σ-algebra can be seen as a collection of all information of the process before and at the stopping time τ.

Definition 2.5. *The process X is said to have the strong Markov property at a finite stopping time τ if for each bounded, measurable function f it holds*

$$\mathbb{E}_x(f(X_{t+\tau})|\mathcal{F}_\tau) = \mathbb{E}_{X_\tau}(f(X_t)) \quad \mathbb{P}_x - a.s.$$

The process X is said to be a strong Markov process if it has the strong Markov property at each finite stopping time τ.

Let $\{\tau_n\}$ be a sequence of stopping times, increasing to τ. A strong Markov process X is said to be quasi-left continuous if $X_{\tau_n} \to X_\tau$ almost surely on $\{\tau < \infty\}$.

Definition 2.6. *A strong Markov process X is said to be a Hunt process if it is right continuous with left limits and quasi-left continuous.*

In many cases we work with a Markov process up to a given stopping time which is then called the lifetime of the process and denoted ζ. Formally, this killed process is defined via

$$\hat{X}_t := \begin{cases} X_t, & t < \zeta, \\ \Delta, & t \geq \zeta, \end{cases}$$

where Δ is a fictitious point and

$$\zeta = \inf\{t : \hat{X}_t = \Delta\}.$$

If X is a strong Markov process then so is $\hat{X}$. By convention, every function f on E equals 0 at Δ.

2.2 One-dimensional diffusions

This section introduces some fundamental facts on real valued diffusion processes. We mainly refer to [13, 26, 32, 46, 47] for further results.

Definition 2.7. *A process X is called a one-dimensional diffusion process (linear diffusion) if it is a time-homogeneous strong Markov process with almost surely continuous sample paths and takes values on an interval $I \subseteq \mathbb{R}$.*

For $y \in I$, denote by

$$H_y := \inf\{t : X_t = y\}$$

the first time the process hits the point y and

$$H_{ab} := H_a \wedge H_b.$$

Definition 2.8. *The diffusion process X is called regular if for every $x, y \in I$*

$$\mathbb{P}_x(H_y < \infty) > 0.$$

This means that the process can visit every point in the state space from every other point. From now on all diffusion processes are assumed to be regular.

Every diffusion process has three basis characteristics which are *the scale function S*, *the speed measure m* and *the killing measure k*. For the present discussion we assume for simplicity that $k \equiv 0$.

• *The scale function:* for $a, b \in I, a < b$ and $x \in [a, b]$, consider the function $q(x) := \mathbb{P}_x(H_b < H_a)$. It is seen that $q(a) = 0, q(b) = 1$.

Proposition 2.9. *There exists a continuous, strictly increasing function S on I such that for all $x \in [a, b]$*

$$q(x) = \mathbb{P}_x(H_b < H_a) = \frac{S(x) - S(a)}{S(b) - S(a)},$$

and

$$\mathbb{P}_x(H_a < H_b) = 1 - \mathbb{P}_x(H_b < H_a) = \frac{S(b) - S(x)}{S(b) - S(a)}.$$

Proof. See Freedman [26, p 113-115], Revuz and Yor [46, p 303]. □

Definition 2.10. *The function S is called the scale function of the diffusion X. If $S(x) \equiv x$ we say that X is on natural scale.*

Notice that if S is a scale function of X then so is $S^*(x) = c_1 + c_2 S(x)$ for any constants c_1 and $c_2 > 0$. Since S is increasing, it is seen that the process $Y_t = S(X_t)$ is a diffusion on the natural scale.

• *The speed measure:* for $a, b \in I, a < b$, define

$$v(x) := \mathbb{E}_x(H_{ab}), \quad a \leq x \leq b.$$

It is seen that $v(a) = v(b) = 0$.

Definition 2.11. *Let $F : [a, b] \to \mathbb{R}$ be a strictly increasing function. A function u is called F-concave if for all $a \leq c < x < d \leq b$*

$$u(x) \geq \frac{F(d) - F(x)}{F(d) - F(c)} u(c) + \frac{F(x) - F(c)}{F(d) - F(c)} u(d).$$

If $-u$ is F-concave we say that u is F-convex. Moreover, the one-sided F-derivatives of u are defined by

$$\frac{d^+u}{dF}(x) := \lim_{\delta\downarrow 0} \frac{u(x+\delta)-u(x)}{F(x+\delta)-F(x)} \quad \textit{and} \quad \frac{d^-u}{dF}(x) := \lim_{\delta\downarrow 0} \frac{u(x-\delta)-u(x)}{F(x-\delta)-F(x)}.$$

The following result and its proof can be found in Freedman [26, p 126] and Revuz and Yor [46, p 304].

Theorem 2.12. *v is continuous and strictly S-concave on $[a,b]$.*

Since v is S-concave then the derivative $\frac{d^+v}{dS}(x)$ exists for all $x \in [a,b]$ and is right continuous and decreasing (see [46, p 544]). Therefore, there is a unique measure m on I defined by

$$m(x,y] := \frac{d^+v}{dS}(x) - \frac{d^+v}{dS}(y), \quad \text{for} \quad x<y, \quad x,y \in I.$$

Proposition 2.13. *It holds that*

$$v(x) = \int_a^b G_{ab}(x,y)m(dy),$$

where

$$G_{ab}(x,y) := \begin{cases} \frac{(S(x)-S(a))(S(b)-S(y))}{S(b)-S(a)}, & a \le x \le y \le b, \\ \\ \frac{(S(y)-S(a))(S(b)-S(x))}{S(b)-S(a)}, & a \le y \le x \le b. \end{cases}$$

Proof. (cf. [46, p 546]). Using the integration by parts and noticing that $v(a) = v(b) = 0$ we have

$$\begin{aligned} \int_a^b G_{ab}(x,y)m(dy) &= \frac{S(b)-S(x)}{S(b)-S(a)} \int_{(a,x]} (S(y)-S(a))m(dy) \\ &\quad + \frac{S(x)-S(a)}{S(b)-S(a)} \int_{(x,b)} (S(b)-S(y))m(dy) \\ &= -\frac{S(b)-S(x)}{S(b)-S(a)} \Big((S(x)-S(a))\frac{d^+v}{dS}(x) - \int_{(a,x]} \frac{d^+v}{dS}(y)dS(y) \Big) \\ &\quad + \frac{S(x)-S(a)}{S(b)-S(a)} \Big((S(b)-S(x))\frac{d^+v}{dS}(x) - \int_{(x,b)} \frac{d^+v}{dS}(y)dS(y) \Big) \\ &= v(x). \end{aligned}$$

□

Definition 2.14. *The measure m in Proposition 2.13 is called the speed measure of the diffusion X.*

According to Itô and McKean[28, p 149], the transition function of a linear diffusion is absolutely continuous with respect to the speed measure, i.e., there is a transition density $p(t, x, y), x, y \in I, t \geq 0$ such that

$$P_t(x, A) = \int_A p(t, x, y) m(dy). \tag{2}$$

Furthermore, $p(t, x, y), t \geq 0$, is jointly continuous in all variables and symmetric, i.e., $p(t, x, y) = p(t, y, x)$.

Definition 2.15. *The infinitesimal generator $\mathcal{A}$ of X is defined via*

$$\mathcal{A}f := \lim_{t \downarrow 0} \frac{P_t f - f}{t}$$

for every continuous, bounded function f on I such that the limit exists in the supremum norm $|| \cdot ||$ and for which it holds

$$\sup_{t>0} \left\| \frac{P_t f - f}{t} \right\| < \infty.$$

We now consider a special case in which the basic characteristics of a diffusion are absolutely continuous with respect to the Lebesgue measure and have smooth derivatives, i.e.,

$$m(dx) = m(x)dx; \quad S(x) = \int^x S'(y)dy.$$

Furthermore, assume that S'' is continuous. Then the infinitesimal generator $\mathcal{A}$ can be written in the form of the second order differential operator (see [13])

$$\mathcal{A}f(x) = \frac{1}{2} a(x) \frac{d^2}{dx^2} f(x) + b(x) \frac{d}{dx} f(x),$$

where the functions a and b (the infinitesimal parameters of X) are related to m and S via

$$m(x) = \frac{2e^{B(x)}}{a^2(x)}, \quad S'(x) = e^{-B(x)}, \quad B(y) = \int^y \frac{b(y)}{a(x)} dx.$$

Let T be an exponentially distributed random variable with parameter $\alpha > 0$, independent of X. Killing X at the time T we obtain the killed process $\hat{X}$ and

$$\mathbb{P}_x(\hat{X}_t \in dy) = \mathbb{P}(X_t \in dy, t < T) = e^{-\alpha t} p(t, x, y) m(dy).$$

The last equality is implied by the independence between T and X. The infinitesimal generator of $\hat{X}$ is of the form

$$\hat{\mathcal{A}} f(x) = \mathcal{A} f(x) - \alpha f(x).$$

Consider the Sturm-Liouville ordinary differential equation

$$\frac{1}{2} a(x) \frac{d^2}{dx^2} f(x) + b(x) \frac{d}{dx} f(x) - \alpha f(x) = 0. \tag{3}$$

It holds that (3) has two linearly independent, continuous, positive solutions ψ_α and φ_α where ψ_α is increasing and φ_α is decreasing. These solutions are unique up to multiplicative constants when the boundary conditions are taken into account. We have that the so-called Wronskian determinant

$$\begin{aligned} w_\alpha := \varphi_\alpha(x) \frac{d^+ \psi_\alpha}{dS}(x) - \psi_\alpha(x) \frac{d^+ \varphi_\alpha}{dS}(x) \\ = \varphi_\alpha(x) \frac{d^- \psi_\alpha}{dS}(x) - \psi_\alpha(x) \frac{d^- \varphi_\alpha}{dS}(x) \end{aligned}$$

is independent of x.

Denote by $G_\alpha(x, y)$ the density of the α-resolvent operator with respect to the speed measure also called the *Green kernel* of X. Then

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

where the transition density $p(t, x, y)$ is with respect to the speed measure, see (2).

Theorem 2.16. *The Green kernel G_α is represented in terms of the functions ψ_α and φ_α as follows*

$$G_\alpha(x, y) = \begin{cases} w^{-1} \psi_\alpha(x) \varphi_\alpha(y), & x \leq y, \\ w^{-1} \varphi_\alpha(x) \psi_\alpha(y), & x \geq y. \end{cases} \tag{4}$$

The proof of Theorem 2.16 can be found in [36, p 31].

Theorem 2.17. *The Laplace transform of the hitting time H_y is given by*

$$\mathbb{E}_x(e^{-\alpha H_y}) = \frac{G_\alpha(x,y)}{G_\alpha(y,y)} = \begin{cases} \frac{\psi_\alpha(x)}{\psi_\alpha(y)}, & x \le y, \\ \frac{\varphi_\alpha(x)}{\varphi_\alpha(y)}, & x \ge y. \end{cases}$$

Proof. See Itô and McKean [28, p. 128]. □

2.3 Lévy processes

This section presents briefly some facts from the theory of stochastic processes with stationary independent increments. These processes are called Lévy processes. Detailed material on Lévy processes can be found, e.g., [6, 10, 29, 34, 49]. We start with the following definition in [34, p 2].

Definition 2.18. *Let $X = (X_t)_{t\ge 0}$ be an adapted stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$. X is said to be a Lévy process if*

(i) $X_0 = 0$ a.s.

(ii) For $0 \le s \le t$, the increment $X_t - X_s$ is independent of $\mathcal{F}_s$,

(iii) For $0 \le s \le t$, $X_t - X_s$ and X_{t-s} have the same law,

(iv) The paths of X are right continuous with left limits a.s.

Some familiar Lévy processes are Brownian motion, compound Poisson processes and Gamma processes. A Lévy process is called a *subordinator* if its paths are non-decreasing almost surely.

Definition 2.19. *A real-valued random variable ξ is said to be infinitely divisible if for any integer n there exists a sequence of i.i.d random variables $\{\xi_i^{(n)}, i \le n\}$ such that*

$$\xi \overset{(d)}{=} \xi_1^{(n)} + \cdots + \xi_n^{(n)}. \tag{5}$$

For $u \in \mathbb{R}$, denote by $\phi(u) := \mathbb{E}(e^{iu\xi})$ the characteristic function of ξ and $\phi_n(u) := \mathbb{E}(e^{iu\xi_1^{(n)}})$ the characteristic function of $\xi_1^{(n)}$. Then (5) is equivalent to

$$\phi(u) = (\phi_n(u))^n.$$

We state next the Lévy-Khintchine representation for infinitely divisible random variables.

Theorem 2.20. *If ξ is an infinitely divisible random variable then there exists a unique triple $(\tilde{a}, \tilde{b}, \nu)$ where $\tilde{a} \in \mathbb{R}$, $\tilde{b} \in \mathbb{R}$ and ν is a measure concentrated on $\mathbb{R} \setminus \{0\}$ satisfying $\int_{\mathbb{R}}(1 \wedge x^2)\nu(dx) < \infty$ such that the characteristic function ϕ of ξ has the representation*

$$\phi(u) = \exp\left\{iu\tilde{a} - \frac{1}{2}\tilde{b}^2u^2 + \int_{\mathbb{R}} \left(e^{iux} - 1 - iux1_{\{|x|<1\}}\right)\nu(dx)\right\}, \quad u \in \mathbb{R}.$$

The measure ν is called the Lévy measure.

By the definition of Lévy processes it is seen that for $t > 0$, the random variable X_t is infinitely divisible. Indeed, for $n = 1, 2, \ldots,$ we may write $X_t = \sum_{k=1}^{n}(X_{kt/n} - X_{(k-1)t/n})$. By the definition of X the random variables $X_{kt/n} - X_{(k-1)t/n}, k = 1, 2, \ldots, n$ are i.i.d. Denote by Ψ the characteristic exponent of X_1, i.e., $\Psi(u) = -\log \mathbb{E}(e^{iuX_1})$ and $\Psi_t(u) := -\log \mathbb{E}(e^{iuX_t})$ then

$$\Psi_t(u) = t\Psi(u).$$

Therefore, the characteristic exponent Ψ of X_1 is called the characteristic exponent of the Lévy process X. Similarly, we also have the Lévy-Khintchine formula for Lévy processes.

Theorem 2.21. *For each $u \in \mathbb{R}$ the characteristic exponent Ψ has the form*

$$\Psi(u) = -iua + \frac{1}{2}b^2u^2 - \int_{\mathbb{R}} \left(e^{iux} - 1 - iux1_{\{|x|<1\}}\right)\nu(dx),$$

where $a, b \in \mathbb{R}$ and ν is the Lévy measure associated with the infinitely divisible random variable X_1.

Let T be an exponentially distributed random variable with parameter $\alpha > 0$, independent of X. For each $u \in \mathbb{R}$ we have

$$\mathbb{E}(e^{iuX_T}) = \frac{\alpha}{\alpha + \Psi(u)}. \tag{6}$$

Indeed,

$$\begin{aligned}\mathbb{E}(e^{iuX_T}) = \int_0^\infty \alpha\ \mathbb{E}(e^{iuX_t})dt &= \alpha \int_0^\infty e^{-(\alpha+\Psi(u))t}dt \\ &= \frac{\alpha}{\alpha + \Psi(u)}.\end{aligned}$$

Let us denote

$$M_t := \sup_{s \le t} X_s \quad \text{and} \quad I_t := \inf_{s \le t} X_s.$$

The processes $(M_t)_{t \ge 0}$ and $(I_t)_{t \ge 0}$ are called the running maximum and running minimum processes associated with X, respectively. It is seen that M_t and $-I_t$ are non-negative increasing, right-continuous and $\mathcal{F}_t$ measurable. The following result is known as the Wiener-Hopf factorization (see, e.g., [29, p 628]).

Theorem 2.22. *The random variables M_T and $X_T - M_T$ are independent and*

$$\mathbb{E}(e^{iuM_T})\mathbb{E}(e^{iu(X_T - M_T)}) = \frac{\alpha}{\alpha + \Psi(u)}. \tag{7}$$

Moreover, I_T and $X_T - M_T$ have the same law.

From (6) and (7) we have the decomposition of X_T as follows

$$X_T \overset{(d)}{=} M_T + I_T^o,$$

where I_T^o is independent of M_T and has the same law as I_T. In practice, we usually use this form of the Wiener-Hopf factorization.

If the Lévy measure ν is supported on $(-\infty, 0)$ we say that X is spectrally negative, equivalently, the process has only negative jumps. In this case it holds that $\mathbb{E}(e^{\lambda X_1}) < \infty$ for all $\lambda > 0$ and, hence, the Laplace exponent of X_1 is given by (see [29, p 632])

$$\begin{aligned}
\Phi(\lambda) :&= \log \mathbb{E}_x(e^{\lambda X_1}) \\
&= -\Psi(-i\lambda) \\
&= \lambda a + \frac{1}{2}\lambda^2 b^2 + \int_{(-\infty,0)} \Big(e^{\lambda x} - 1 - \lambda x 1_{\{-1<x<0\}}\Big)\nu(dx).
\end{aligned}$$

If ν is supported on $(0, +\infty)$ we say that X is spectrally positive or the process has only positive jumps. Similarly, we have $\mathbb{E}(e^{-\lambda X_1}) < \infty$ for all $\lambda > 0$, i.e., X_1 has the Laplace exponent. For spectrally one-sided Lévy processes we have the following result.

Proposition 2.23.

(i) If X is spectrally negative then M_T is exponentially distributed with mean $1/q(\alpha)$, where $q(\alpha)$ is the unique positive root of $\Phi(\lambda) = \alpha$, $\lambda > 0$.

(ii) If X is spectrally positive then $-I_T$ is exponentially distributed with mean $1/\hat{q}(\alpha)$, where $\hat{q}(\alpha)$ is the unique positive root of $\Phi(-\lambda) = \alpha$, $\lambda > 0$.

Proof. (cf. [29, p 632]). We prove case (i), case (ii) can be carried out similarly. For the uniqueness of the positive root, see Kyprianou [34, p 81]. For $y > 0$ let $\tau(y) := \inf\{t : X_t \geq y\}$. Notice that for all $\lambda > 0$ the process $(e^{\lambda X_t - t\Phi(\lambda)})_{t\geq 0}$ is a martingale. Now applying the optional sampling theorem at the stopping time $\tau(y) \wedge t$ and replacing λ by $q(\alpha)$ we obtain

$$\mathbb{E}(e^{q(\alpha)X_{\tau(y)\wedge t} - (\tau(y)\wedge t)\alpha}) = 1. \tag{8}$$

Because of the absence of positive jumps we have $X_{\tau(y)} = y$ on $\{\tau(y) < \infty\}$ and, hence, $q(\alpha)X_{\tau(y)\wedge t} - (\tau(y) \wedge t)\lambda < q(\alpha)y$. In (8) let t tend to $+\infty$ to obtain

$$\mathbb{E}(e^{-\alpha\tau(y)} 1_{\{\tau(y)<\infty\}}) = e^{-yq(\alpha)}.$$

So for any $y > 0$ we have

$$\begin{aligned}\mathbb{P}(M_T \geq y) = \mathbb{P}(\tau(y) \leq T) &= \mathbb{E}\Big(\int_{\tau(y)}^{\infty} \alpha\, e^{-\alpha s} ds\Big) \\ &= \mathbb{E}(e^{-\alpha\tau(y)} 1_{\{\tau(y)<\infty\}}) = e^{-yq(\alpha)}.\end{aligned}$$

□

3 Excessive functions

The theory of excessive functions plays a crucial role not only in the theory of Markov processes, but also in the theory of optimal stopping. It will be showed in section 5 that the value function of an optimal stopping problem is, in fact, α-excessive. We present here some basic facts on excessive functions for Markov processes. We refer to [11, 13, 17, 21] for more detailed results on this topic.

3.1 Basic properties

Definition 3.1. *Let $\alpha \geq 0$. A measurable function $f : E \mapsto \mathbb{R}_+ \cup \{+\infty\}$ is said to be α-excessive for a real-valued Markov process X if*

(a) $e^{-\alpha t}\mathbb{E}_x(f(X_t)) \leq f(x)$ *for all* $x \in E, t \geq 0$,

(b) $\lim_{t\downarrow 0} e^{-\alpha t}\mathbb{E}_x(f(X_t)) = f(x)$ *for all* $x \in E$.

The following results and their proofs can be found in [11, 17].

Proposition 3.2.

(i) Let f and g be α-excessive functions and c a non-negative constant. Then the functions $f+g$, $f \wedge g$ and cf are α-excessive.

(ii) If $\{f_n\}$, $n = 1, 2, \ldots$ is an increasing sequence of α-excessive functions then $f^* := \lim_n f_n$ is α-excessive.

Proposition 3.3. *If f is α-excessive then there exists functions g_n, $n = 1, 2, \ldots$ such that $U^\alpha g_n$ increases pointwise to f as n tends to ∞.*

Next we recall the martingale characterization of the excessive functions.

Proposition 3.4. *If f is α-excessive and finite then $(e^{-\alpha t} f(X_t))_{t \geq 0}$ is a non-negative supermartingale.*

Proof. (cf. [17]). Using the Markov property we have for $t \geq s$

$$\mathbb{E}_x(e^{-\alpha t} f(X_t) | \mathcal{F}_s) = e^{-\alpha s} \mathbb{E}_{X_s}(e^{-\alpha(t-s)} f(X_{t-s})) \leq e^{-\alpha s} f(X_s).$$

Furthermore, $\mathbb{E}_x |e^{-\alpha t} f(X_t)| = \mathbb{E}_x(e^{-\alpha t} f(X_t)) \leq f(x) < \infty$. □

Corollary 3.5. *Let f be an α-excessive function and finite.*
(i) For finite stopping times τ_1 and τ_2 such that $\tau_1 \leq \tau_2$ it holds

$$\mathbb{E}_x(e^{-\alpha \tau_2} f(X_{\tau_2})) \leq \mathbb{E}_x(e^{-\alpha \tau_1} f(X_{\tau_1})).$$

In particular, for $\tau_1 \equiv 0$

$$\mathbb{E}_x(e^{-\alpha \tau_2} f(X_{\tau_2})) \leq f(x).$$

(ii) If τ is a finite stopping time and $g(x) := \mathbb{E}_x(e^{-\alpha \tau} f(X_\tau))$ then g is α-excessive.

For a non-negative and lower semicontinuous function f, it is proved in Theorem 12.4(A) in [21] that f is α-excessive if and only if

$$\mathbb{E}_x(e^{-\alpha \tau_D} f(X_{\tau_D})) \leq f(x)$$

for all x and for all $\tau_D = \inf\{t : X_t \in D\}$, where D is an arbitrary compact subset of E and we define $f(X_{\tau_D}) = 0$ if $\tau_D = \infty$. When applying this fact for a regular diffusion X we obtain the following useful result.

Proposition 3.6. *Let X be a regular diffusion. Then the fundamental solutions ψ_α and φ_α are α-excessive functions. Consequently, the Green kernel $G_\alpha(x,y)$ is α-excessive for each $y \in I$.*

Proof. Since the sample paths of X are continuous we may, without loss of generality, take $D = \{y\}$, where y is an arbitrary point of I. Then $\tau_D = H_y$ and we have

$$\mathbb{E}_x(e^{-\alpha H_y}\psi_\alpha(X_{H_y})) = \psi_\alpha(y)\mathbb{E}_x(e^{-\alpha H_y}).$$

Combining with the fact

$$\mathbb{E}_x(e^{-\alpha H_y}) = \begin{cases} \frac{\psi_\alpha(x)}{\psi_\alpha(y)}, & x \leq y, \\ \frac{\varphi_\alpha(x)}{\varphi_\alpha(y)}, & x \geq y, \end{cases}$$

we get

$$\mathbb{E}_x(e^{-\alpha H_y}\psi_\alpha(X_{H_y})) = \begin{cases} \psi_\alpha(x), & x \leq y, \\ \psi_\alpha(y)\frac{\varphi_\alpha(x)}{\varphi_\alpha(y)}, & x \geq y. \end{cases}$$

Since φ_α and ψ_α are decreasing and increasing, respectively, the function $x \mapsto \frac{\varphi_\alpha(x)}{\psi_\alpha(x)}$ is decreasing. Hence, for $x \geq y$ we have $\psi_\alpha(y)\frac{\varphi_\alpha(x)}{\varphi_\alpha(y)} \leq \psi_\alpha(x)$. So for all $x \in I$ we obtain

$$\mathbb{E}_x(e^{-\alpha H_y}\psi_\alpha(X_{H_y})) \leq \psi_\alpha(x).$$

The proof for φ_α is carried out similarly. Since the function $x \mapsto G_\alpha(x,y)$ can be represented as the minimum of the functions ψ_α and φ_α it follows from Proposition 3.2 that it is α-excessive. □

3.2 Representation theory

We discuss next the Riesz and Martin integral representations of excessive functions. We focus on linear diffusions and extract the basic theorems from Article II. For results for general continuous time Markov processes, e.g., for Hunt processes, see [17] and [33]. These presentations can be used when analyzing and solving optimal stopping problems, see [4, 16, 15, 18, 38, 48].

Let X be a linear diffusion and recall that X has an α-resolvent kernel G_α given by (4). In particular, G_α is symmetric, i.e., $G_\alpha(x,y) = G_\alpha(y,x)$. This implies that X is self dual, i.e., for all Borel subsets A and B of I it holds

$$\int_B G_\alpha(x,A)m(dx) = \int_A G_\alpha(y,B)m(dy).$$

Due to the self duality and the regularity properties of the resolvent kernel we have the following result.

Theorem 3.7. (Riesz representation). *Each α-excessive function f of a linear diffusion X can be decomposed uniquely in the form*

$$f(x) = \int_I G_\alpha(x,y)\sigma(dy) + h_\alpha(x), \tag{9}$$

where h_α is an α-harmonic function and σ is a Radon measure on $\mathbb{R}$. Moreover, the measure σ is unique.

Here, the α-harmonicity of h_α means that for all compact subset K of I it holds

$$\mathbb{E}_x(e^{-\alpha\tau_K} h_\alpha(X_{\tau_K})) = h_\alpha(x),$$

where

$$\tau_K := \inf\{t : X_t \notin K\}.$$

The Martin representation of an α-excessive function gives also the information of the α-harmonic function h_α in the Riesz representation. Next theorem gives an explicit form of the Martin representation (see [13, p 33], [48]).

Theorem 3.8. *Let X be a linear diffusion and f a function on I such that $f(x_0) = 1$ for some $x_0 \in I$. Then f is α-excessive if and only if there exists a probability measure ν_f on I such that*

$$f(x) = \int_{(l,r)} \frac{G_\alpha(x,y)}{G_\alpha(x_0,y)}\nu_f(dy) + \frac{\varphi_\alpha(x)}{\varphi_\alpha(x_0)}\nu_f(\{l\}) + \frac{\psi_\alpha(x)}{\psi_\alpha(x_0)}\nu_f(\{r\}), \tag{10}$$

where l and r are the left and right end points of I, respectively.

Remark 3.9. *It is proved in [4] that α-harmonic function h_α has the form*

$$h_\alpha(x) = \begin{cases} c_1\psi_\alpha(x), & \text{if } \ l \in I, r \notin I, \\ c_2\varphi_\alpha(x), & \text{if } \ l \notin I, r \in I, \\ c_1\psi_\alpha(x) + c_2\varphi_\alpha(x), & \text{if } \ l \notin I, r \notin I. \end{cases}$$

The representing measure ν_f in the Martin representation is given as follows (see [48]).

Proposition 3.10. *Let f be an α-excessive function for a regular diffusion and x_0 a point such that $f(x_0) = 1$. Then the representing measure ν_f of f is given by*

$$\nu_f((x,r]) = \frac{\psi_\alpha(x_0)}{\omega_\alpha}(\varphi_\alpha(x)f^+(x) - f(x)\varphi_\alpha^+(x)), \quad x \geq x_0,$$

and

$$\nu_f([l,x)) = \frac{\varphi_\alpha(x_0)}{\omega_\alpha}(f(x)\psi_\alpha^-(x) - \psi_\alpha(x)f^-(x)), \quad x \leq x_0,$$

where the superscripts $^+$ and $^-$ denote the right and left derivatives with respect to the scale function S, respectively.

Remark 3.11. *It is seen that the representing measure σ in the Riesz representation and the representing measure ν_f in the Martin representation for diffusions are connected inside I via*

$$\sigma(dy) = \frac{\nu_f(dy)}{G_\alpha(x_0,y)}.$$

The next result is also from [48] and is essential when studying the validity of the principle of smooth fit in optimal stopping problems for diffusions.

Proposition 3.12. *Let X be a diffusion taking values in I. Assume that for a inner point x of I the derivatives $\frac{d\varphi_\alpha}{dS}(x)$ and $\frac{d\psi_\alpha}{dS}(x)$ exist. Then for every α-excessive function f*

$$\frac{d^-f}{dS}(x) \geq \frac{d^+f}{dS}(x)$$

and $\frac{df}{dS}(x)$ exists if and only if $\sigma\{x\} = 0$.

Remark 3.13. *Convexity properties of excessive functions have been studied, e.g., in [3, 19]. In particular, for Brownian motion killed at 0 and 1 it holds that a function is excessive if and only if it is concave (see Dynkin and Yushkevich [22, p 119]). The extended result to an arbitrary diffusion X is stated in Proposition 3.1 in [19]. Another result for Brownian motion is that every α-excessive function which is α-harmonic on a subinterval of $\mathbb{R}$ is, in fact, convex on this interval.*

4 Appell polynomials

In Article III we introduce Appell polynomials associated with a random variable. Here, we present some results for more general Appell polynomials.

4.1 Basic definitions

Definition 4.1. *For $x \in \mathbb{R}$, polynomials $Q_n(x), n = 0, 1, 2, \ldots$ of order n with $Q_0 \equiv c$ (some non-zero real constant) satisfying the condition*

$$\frac{d}{dx} Q_n(x) = n Q_{n-1}(x) \tag{11}$$

are called Appell polynomials.

Such polynomials were first introduced by Appell in his paper [5] published in 1880. It is seen that a polynomial Q_n is an Appell polynomial if and only if it has the form

$$Q_n(x) = \sum_{k=0}^{n} \binom{n}{k} Q_k(0) x^{n-k}. \tag{12}$$

Indeed, we can obtain this from (11) by using induction. The converse is clear. The representation (12) shows that the sequence $\{Q_n(0)\}_{n=0}^{\infty}$ determines the whole family $\{Q_n\}, n = 0, 1, \ldots$

Remark 4.2. *Given a random variable ξ with moments of all orders. The Appell polynomials associated ξ, say, $Q_n^{(\xi)}$ can be defined via the recursive differential equation (11) together with the normalisation*

$$\mathbb{E}(Q_n^{(\xi)}(\xi)) = 0.$$

Next we consider the approach to Appell polynomials via generating functions (see [45],[52]). For a given real analytic function f consider the expansion

$$f(t) e^{xt} = \sum_{n=0}^{\infty} \frac{t^n}{n!} P_n(x), \tag{13}$$

where P_n is a polynomial of order n. Straightforward calculations show then that $\{P_n\}$ satisfies the differential equation (11), i.e., the polynomials P_n are, in fact, Appell polynomials. The function f is called the generating function or the determining function of the Appell polynomials $P_n, n = 0, 1, 2, \ldots$ Notice also that from (13) it follows that

$$f(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} P_n(0).$$

Now given a random variable ξ with finite exponential moments it is seen that the function $t \mapsto (\mathbb{E}(e^{t\xi}))^{-1}$ is analytic. Suggested by (13), we state now the definition of the Appell polynomials associated with ξ.

Definition 4.3. *Let ξ be a random variable having some exponential moments. Polynomials $Q_n^{(\xi)}$, $n = 0, 1, 2, \ldots$ being of order n, satisfying*

$$\frac{e^{xt}}{\mathbb{E}(e^{t\xi})} = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q_n^{(\xi)}(x)$$

are called the Appell polynomials associated with ξ.

These polynomials have some important applications in probability theory, e.g., in optimal stopping as was first noticed by Novikov and Shiryaev [39].

Remark 4.4. *Sheffer [51] studied a wider class of polynomials containing Appell polynomials (see also [45, p 222]) which are generated from*

$$f(t)e^{g(t)x} = \sum_{n=0}^{\infty} \frac{t^n}{n!} S_n(x),$$

where f and g are two analytic functions satisfying $g(0) = 0, g'(0) \neq 0$ and $f(0) \neq 0$. The polynomials $S_n(x)$, $n = 1, 2, \ldots$ are called the Sheffer polynomials. These polynomials can be applied to construct, e.g., time-space martingales for Lévy processes, for further details we refer to Schoutens [50].

4.2 Integral representation

As noticed above the Appell polynomials $Q_n, n = 0, 1, \ldots$ are completely determined by the sequence $\{Q_n(0)\}_{n=0}^{\infty}$ this allows us to characterize the Appell polynomials via so-called moment problems. For this we recall first a famous result of Boas [12], see also Widder [58, p 139].

Theorem 4.5. *Let $\{\mu_n\}$ be an arbitrary sequence of real numbers. Then there exists a function of bounded variation, say, γ such that*

$$\mu_n = \int_0^{\infty} t^n d\gamma(t), \quad n = 0, 1, 2, \ldots$$

Utilizing this result, Sheffer [52] derived the following characterization of Appell polynomials which we wish to present here since it is not appearing in Article III.

Theorem 4.6. *Polynomials $Q_n, n = 1, 2, \ldots,$ are Appell polynomials if and only if there exists a function γ of bounded variation on $(0, \infty)$ such that for $n = 0, 1, \ldots$*

(i) the constants $c_n := \int_0^\infty t^n d\gamma(t)$ exist and are finite,

(ii) $c_0 \neq 0$,

(iii) $Q_n(x) = \int_0^\infty (x+t)^n d\gamma(t)$.

Proof. Assume that Q_n is an Appell polynomial. Then

$$Q_n(x) = \sum_{k=0}^{n} \binom{n}{k} Q_k(0) x^{n-k}.$$

From Theorem 4.5 the sequence $\{Q_n(0)\}_{n=0}^\infty$ can be represented as

$$Q_n(0) = \int_0^\infty t^n d\gamma(t),$$

where γ is of bounded variation. Consequently,

$$Q_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} \int_0^\infty t^k d\gamma(t) = \int_0^\infty (x+t)^n d\gamma(t).$$

For the necessary condition it is easily checked that the polynomial Q_n in (iii) satisfies differential equation (11). □

From this theorem it is seen that if γ is a probability distribution of some positive random variable η then the Appell polynomials Q_n has the moment representation

$$Q_n(x) = \mathbb{E}(x+\eta)^n, \quad n = 0, 1, 2, \ldots$$

Let F be the probability distribution of a random variable ζ, not necessarily positive, but assumed to have the moments of all orders. Then the polynomials

$$\hat{Q}_n(x) := \int_{\mathbb{R}} (x+t)^n dF(t) = \mathbb{E}(x+\zeta)^n \tag{14}$$

satisfy (11) and are, in fact, Appell polynomials. We now present a condition such that polynomials P_n as defined in (13) admit representation (14). The following result is known as the solution to the Hamburger moment problem (see [58, p 129-135]).

Theorem 4.7. *The necessary and sufficient condition for a given sequence* $\{\mu_n\}$ *to be represented as*

$$\mu_n = \int_{\mathbb{R}} t^n dF(t),$$

where F *is a probability distribution is that the quadratic form*

$$\sum_{i=0}^{n} \sum_{j=0}^{n} \mu_{i+j} p_i p_j \tag{15}$$

is, for all $n \geq 0$*, positive definite or equivalently the matrix* $A_n = (\mu_{i+j})_{ij}, i, j = 0, 1, \ldots, n$ *is for all* $n \geq 0$ *positive definite.*

Corollary 4.8. *Appell polynomials* $P_n, n = 0, 1, 2, \ldots$ *given via (13) admit representation (14) if and only if the sequence* $\{P_n(0)\}$ *induces the positive definite quadratic forms.*

In Article III a simple proof is given that there does not exist a non-trivial real random variable ζ such that the Appell polynomial $Q_n^{(\xi)}$ associated with a random variable ξ having some exponential moments can be represented as a n-th moment of $x + \zeta$. We provide here another explanation of this fact. Putting $h(t) := (\mathbb{E}(e^{t\xi}))^{-1}$ we get the Taylor series

$$h(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} h^{(n)}(0).$$

Clearly, $\mu_0 := h^{(0)}(0) = 1$ and

$$\mu_1 := h^{(1)}(0) = -\mathbb{E}(\xi), \quad \mu_2 := h^{(2)}(0) = 2(\mathbb{E}(\xi))^2 - \mathbb{E}(\xi^2).$$

Hence, the determinant

$$\begin{vmatrix} \mu_0 & \mu_1 \\ \mu_1 & \mu_2 \end{vmatrix} = (\mathbb{E}(\xi))^2 - \mathbb{E}(\xi^2) \leq 0,$$

which implies that the quadratic form (15) is not for $n = 2$ positive definite. So the sequence $\{h^{(n)}(0)\}$ is not a moment sequence of a probability distribution, i.e., $Q_n^{(\xi)}$ does not have representation (14).

5 Optimal stopping theory

In this section we present some fundamental results from the theory of optimal stopping. We refer to Shiryaev [53] and Peskir and Shiryaev [42] and the references therein.

5.1 Excessive characterization of the value function

Let $X = (X_t)_{t\geq 0}$ be a continuous time real valued strong Markov process starting at $x \in \mathbb{R}$. Consider the optimal stopping problem of finding a value function V and a stopping time τ^* such that

$$V(x) := \sup_{\tau\in\mathcal{M}} \mathbb{E}_x(e^{-\alpha\tau}G(X_\tau)) = \mathbb{E}_x(e^{-\alpha\tau^*}G(X_{\tau^*})), \tag{16}$$

where $\mathcal{M}$ denotes the set of all stopping times with respect to the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ of X, G is a non-negative, continuous function and the discounting factor $\alpha \geq 0$ such that

$$\mathbb{E}_x(\sup_{t\geq 0} e^{-\alpha t}G(X_t)) < \infty. \tag{17}$$

In case $\tau = +\infty$ we set

$$e^{-\alpha\tau}G(X_\tau) = \limsup_{t\to\infty} e^{-\alpha t}G(X_t).$$

We now present some general results from the theory of optimal stopping. We start with the definition of the notion "excessive majorant" which is of the key importance in this theory.

Definition 5.1. *Let g be a non-negative function. A function f is called an α-excessive majorant of g if f is α-excessive and $f(x) \geq g(x)$ for all x.*
A function f is called the smallest α-excessive majorant of g if

(i) f is an α-excessive majorant of g.

(ii) Every α-excessive majorant $\hat{f}$ of g dominates f.

The next result is due to Grigelionis and Shiryaev [27] (see also [9, 40], [53, chapter 3]), it provides a method to construct the smallest α-excessive majorant from a given reward function G.

Theorem 5.2. *Assume that the reward function G is non-negative and continuous. Let $S_n = \{k2^{-n} : 0 \leq k \leq n2^n\}$, and define*

$$v_0(x) := G(x), \quad v_n(x) := \sup_{t \in S_n} \mathbb{E}_x(e^{-\alpha t} v_{n-1}(X_t))$$

for $n = 1, 2, \ldots$ Then the sequence $\{v_n\}$ is increasing and the function $v := \lim_{n\to\infty} v_n$ is the smallest α-excessive majorant of G. Moreover, v is lower semi-continuous.

The excessive characterization of the value function is stated in the following theorem (see [20]).

Theorem 5.3. *Suppose that the reward G is non-negative, continuous and satisfies (17). Then the value function V is the smallest α-excessive majorant of G. Moreover, an optimal stopping time is given by $\tau^* := \inf\{t : X_t \in \Gamma\}$, where $\Gamma := \{x : V(x) = G(x)\}$ is the so called stopping region.*

Proof. (cf. [9, 40, 42]) We prove the result under the conditional simplifying assumptions that τ^* is finite, a.s. By Theorem 5.2 the function v is defined therein the smallest α-excessive majorant of G and lower semicontinuous. Denote $C := \{x : G(x) < v(x)\}$ and $\hat{\tau}^* := \inf\{t \geq 0 : X_t \notin C\}$. For $\epsilon > 0$, setting $C_\epsilon := \{x \in E : G(x) < v(x) - \epsilon\}$ and $\tau_\epsilon := \inf\{t : X_t \notin C_\epsilon\}$, it is seen that $C_\epsilon \subseteq C$ and $\tau_\epsilon \leq \hat{\tau}^*$. From Corollary 3.5 we know that the function

$$\hat{V}_\epsilon(x) := \mathbb{E}_x(e^{-\alpha\tau_\epsilon} v(X_{\tau_\epsilon}))$$

is α-excessive. We claim that for all $x \in E$

$$G(x) \leq \hat{V}_\epsilon(x) + \epsilon. \tag{18}$$

Indeed, assume that (18) does not hold. Then

$$\lambda := \sup_x (G(x) - \hat{V}_\epsilon(x)) > \epsilon. \tag{19}$$

It is seen from (19) that the function $\hat{V}_\epsilon + \lambda$ is α-excessive majorant of G. Hence, for all x

$$v(x) \leq \hat{V}_\epsilon(x) + \lambda. \tag{20}$$

Choosing $\eta < \epsilon$ and x_0 such that

$$G(x_0) - \hat{V}_\epsilon(x_0) > \lambda - \eta, \tag{21}$$

We get from (21) and (20)

$$v(x_0) \leq \hat{V}_\epsilon(x_0) + \lambda \leq G(x_0) + \eta. \tag{22}$$

Consequently, we have

$$\begin{aligned} G(x_0) + \eta &\geq v(x_0) \\ &\geq \mathbb{E}_{x_0}(e^{-\alpha(\tau_\epsilon \wedge t)} v(X_{\tau_\epsilon \wedge t})) \\ &\geq \mathbb{E}_{x_0}(e^{-\alpha t}(G(X_t) + \epsilon); t \leq \tau_\epsilon), \end{aligned}$$

where the first inequality is due to (22), the second one is implied by the fact that v is α-excessive and the last one is due to the definition of τ_ϵ. Using the continuity of G and the Fatou's lemma we obtain

$$\begin{aligned} G(x_0) + \eta &\geq \mathbb{E}_{x_0}\Big(\liminf_{t \to 0} \left(e^{-\alpha t}(G(X_t) + \epsilon) 1_{t \leq \tau_\epsilon} \right) \Big) \\ &= G(x_0) + \epsilon \end{aligned}$$

which contradicts the choice of η. Hence, the statement (18) holds. It follows that the function $\hat{V}_\epsilon + \epsilon$ is an α-excessive majorant of G. Therefore, for all x we have

$$\begin{aligned} v(x) \leq \hat{V}_\epsilon(x) + \epsilon &= \mathbb{E}_x(e^{-\alpha\tau_\epsilon} v(X_{\tau_\epsilon})) + \epsilon \\ &\leq \mathbb{E}_x(e^{-\alpha\tau_\epsilon}(G(X_{\tau_\epsilon}) + \epsilon)) + \epsilon \\ &= \mathbb{E}_x(e^{-\alpha\tau_\epsilon} G(X_{\tau_\epsilon}) + 2\epsilon \\ &\leq \sup_\tau \mathbb{E}_x(e^{-\alpha\tau} G(X_\tau) + 2\epsilon = V(x) + 2\epsilon. \end{aligned} \tag{23}$$

Since ϵ is arbitrary it follows that $v(x) \leq V(x)$ for all x. On the other hand, for any stopping time τ by Proposition 3.5 we have

$$v(x) \geq \mathbb{E}_x(e^{-\alpha\tau} v(X_\tau)) \geq \mathbb{E}_x(e^{-\alpha\tau} G(X_\tau)).$$

Taking the supremum over the set $\mathcal{M}$ yields

$$v(x) \geq \sup_\tau \mathbb{E}_x(e^{-\alpha\tau} v(X_\tau)) = V(x).$$

Therefore, we obtain $v(x) = V(x)$ for all x, i.e., V is the smallest α-excessive majorant of G and $\hat{\tau}^* \equiv \tau^*$. Since τ^* is assumed to be finite and G is continuous and $\tau_\epsilon \uparrow \tau^*$ as $\epsilon \downarrow 0$, it follows that $\hat{V}_\epsilon(x) = \mathbb{E}_x(e^{-\alpha\tau_\epsilon} G(X_{\tau_\epsilon})) \to \mathbb{E}_x(e^{-\alpha\tau^*} G(X_{\tau^*}))$. From inequality (23) we have

$$V(x) = v(x) \leq \mathbb{E}_x(e^{-\alpha\tau_\epsilon} G(X_{\tau_\epsilon})) + 2\epsilon \leq V(x) + 2\epsilon.$$

Letting ϵ tend to 0 yields

$$V(x) = \mathbb{E}_x(e^{-\alpha\tau^*} G(X_{\tau^*})),$$

and τ^* is an optimal stopping time. □

Remark 5.4. *(i) The results in Theorem 5.2 and Theorem 5.3 still hold if the condition of continuity of the reward is replaced by the lower semicontinuity. (ii) If G is upper semicontinuous and there exists a lower semicontinuous, α-excessive majorant $\tilde{V}$ of G and $\tau_0 < \infty$ $\mathbb{P}_x$-a.s. for all x, where $\tau_0 = \inf\{t : X_t \in \Gamma_0\}$, $\Gamma_0 = \{x : \tilde{V}(x) = G(x)\}$ then $\tilde{V} = V$ and τ_0 is an optimal stopping time (see [42, p 40-41]).*

We also state the following corollary which gives a characterization for an optimal stopping time (see [40, p 213]).

Corollary 5.5. *Assume that there exists a Borel set D and an α-excessive majorant U of the reward function G such that for all x*

$$U(x) = \mathbb{E}_x(e^{-\alpha\tau_D} G(X_{\tau_D})),$$

where $\tau_D = \inf\{t : X_t \notin D\}$. Then $U(x) = V(x)$ for all x and τ_D is an optimal stopping time.

Proof. Since V is the smallest α-excessive majorant of G we have $V(x) \leq U(x)$ for all x. On the other hand,

$$U(x) = \mathbb{E}_x(e^{-\alpha\tau_D} G(X_{\tau_D})) \leq \sup_\tau \mathbb{E}_x(e^{-\alpha\tau} G(X_\tau)) = V(x).$$

So, $U \equiv V$ and τ_D is optimal. □

Proposition 5.6. *Assume that there exists an optimal stopping time $\hat{\tau}$ of OSP (16). Then τ^* as defined in Theorem 5.3 satisfies*

$$\tau^* \leq \hat{\tau}.$$

Proof. (cf. [40, 56]) Since V is an α-excessive majorant of G and $\hat{\tau}$ is an optimal stopping time we have

$$V(x) = \mathbb{E}_x(e^{-\alpha\hat{\tau}}G(X_{\hat{\tau}})) \leq \mathbb{E}_x(e^{-\alpha\hat{\tau}}V(X_{\hat{\tau}})) \leq V(x).$$

Hence

$$V(x) = \mathbb{E}_x(e^{-\alpha\hat{\tau}}V(X_{\hat{\tau}})). \tag{24}$$

If $\hat{\tau} < \tau^*$ then $G(X_{\hat{\tau}}) < V(X_{\hat{\tau}})$ and, hence,

$$V(x) = \mathbb{E}_x(e^{-\alpha\hat{\tau}}G(X_{\hat{\tau}})) < \mathbb{E}_x(e^{-\alpha\hat{\tau}}V(X_{\hat{\tau}})) = V(x),$$

which is a contradiction proving that $\hat{\tau} \geq \tau^*$. □

5.2 The principle of smooth fit

The principle of smooth fit means that if x^* is a boundary point of the stopping set in OSP (16) then we expect that the value function is continuous and differentiable at x^*, i.e.,

$$\begin{cases} V(x^*) = G(x^*), \\ V'(x^*) = G'(x^*). \end{cases}$$

The principle of smooth fit −sometimes also called the principle of smooth pasting or high contact− is a basic tool to solve optimal stopping problems for one−dimensional diffusions when the reward is differentiable and the underlying process has smooth characteristics. However, in general, the smooth fit may not hold for diffusions with non-smooth characteristics and processes with jumps, e.g., Lévy processes. We refer to [1, 15, 35] for discussions of the smooth fit of concrete optimal stopping problems for Lévy processes.

To indicate how the smooth fit is used to solve OSPs we present here the verification theorem from Øksendal [40, p 224] which, in particular, is applicable for diffusions. The theorem can be used so that we "guess" the form of the continuation set, denoted D in the theorem, and find a function u satisfying, in particular, property (ii).

Theorem 5.7. *Let X be a one-dimensional diffusion with the generator $\mathcal{A}$ and consider OSP (16). Assume that there exists a function $u \geq G$ such that*

(i) $u \in C^2(\mathbb{R} \setminus \partial D)$, where $D := \{x : G(x) < u(x)\}$ and u'' is locally bounded near ∂D,

(ii) $\mathcal{A}u \le 0$ *on* $\mathbb{R} \setminus D$ *and* $\mathcal{A}u = 0$ *on* D*,*

(iii) $\tau_D := \inf\{t : X_t \notin D\} < \infty$ $\mathbb{P}_x$*-a.s for all* x*,*

(iv) for any stopping time τ*, the family of random variables* $\{u(X_\tau), \tau < \tau_D\}$ *is uniformly integrable w.r.t* $\mathbb{P}_x$ *for all* x*.*

Then

$$u(x) = V(x) = \sup_\tau \mathbb{E}_x(e^{-\alpha\tau} G(X_\tau))$$

and τ_D *is an optimal stopping time.*

6 Supplementary examples

In this section we give some new additional examples which highlight the results in Article I. In these examples T denotes an exponentially distributed random variable with parameter $\alpha > 0$, independent of the underlying process X.

6.1 Reflecting Brownian motion

We apply a method which is presented in Article I to study OSP (16) for reflecting Brownian motion and the reward $G(x) = (x^+)^n$, $n = 1, 2, \ldots$ We refer to Theorem 2.5 in Article I for a method of finding the function $\hat{f}$. According to this method we first introduce

$$\tilde{f}_n(x) := (\alpha - \mathcal{A})x^n = x^{n-2}\Big(\alpha x^2 - \frac{n(n-1)}{2}\Big), \quad x \ge 0. \tag{25}$$

where $\mathcal{A} = \frac{1}{2}\frac{d^2}{dx^2}$ is the differential operator associated with X and then investigate whether the equality (16) (see p 1146), i.e.,

$$\frac{1}{\alpha}\mathbb{E}_x(\tilde{f}_n(X_T)) = x^n, \quad x \ge 0 \tag{26}$$

holds.

Remark 6.1. *The polynomials* $\frac{1}{\alpha}\tilde{f}_n$ *in (25) are the Appell polynomials associated with* $|B_T|$*.*

In most cases the function $\tilde{f}$ is a good candidate for finding the needed function $\hat{f}$. However, in this example for $n = 1$ identity (26) does not hold. In fact, we have the following proposition.

Proposition 6.2. *Let $\tilde{f}_n$ be as in (25). Then for all $x \geq 0$*

(a) $\frac{1}{\alpha}\mathbb{E}_x(\tilde{f}_1(|B_T|) = x + \frac{1}{\sqrt{2\alpha}}e^{-x\sqrt{2\alpha}}$.

(b) $\frac{1}{\alpha}\mathbb{E}_x(\tilde{f}_n(|B_T|) = x^n, \quad n \geq 2.$

Proof. It holds

$$\mathbb{E}_x(e^{u|B_T|}) = \frac{\sqrt{2\alpha}}{2\alpha - u^2}\Big(\sqrt{2\alpha}e^{ux} + ue^{-x\sqrt{2\alpha}}\Big), \quad |u| < \sqrt{2\alpha},$$

see formula 1.0.3 in [13, p. 333]. Taking the derivative in u and letting $u = 0$ yields (a).
(b) For simplicity, we consider the case $\alpha = 1/2$. Writing

$$(1 - u^2)\mathbb{E}_x(e^{u|B_T|}) = e^{ux} + ue^{-x},$$

and expanding in the Taylor series yield

$$\sum_{n=0}^{\infty} \frac{1}{n!}(u^n - u^{n+2})\mathbb{E}_x(|B_T|^n) = \sum_{n=0}^{\infty} \frac{u^n}{n!}x^n + ue^{-x}.$$

Identifying herein the coefficients of u^n gives for $n \geq 2$

$$2\mathbb{E}_x(\tilde{f}_n(|B_T|)) = \mathbb{E}_x(|B_T|^{n-2}(|B_T|^2 - n(n-1))) = x^n.$$

□

Remark 6.3. *Since the considered process is not spatially homogeneous the mean value theorem of Appell polynomials cannot be applied to deduce identity (b).*

From identity (b) in Proposition 6.2 it is seen that we can apply for $n \geq 2$ the method in Article I. We study this case first.

Case 1. $n \geq 2$. Since the function $\tilde{f}_n$ satisfies identity (26) it is a candidate to find the function $\hat{f}_n$. Using the formula in Lemma 2.12, p. 1147-Article I (or formulae 1.1.2 and 1.1.6 on p. 333 in [13]) gives

$$\mathbb{P}_x(X_T \in dy | M_T = z) = \sqrt{2\alpha}\ \frac{\cosh(y\sqrt{2\alpha})}{\sinh(z\sqrt{2\alpha})}dy, \quad z > 0. \tag{27}$$

We now consider the function

$$
\begin{aligned}
Q_n(z) &:= \frac{1}{\alpha}\int_0^z \tilde{f}_n(y)\mathbb{P}_x(X_T \in dy | M_T = z) \\
&= \frac{\sqrt{2\alpha}}{\alpha \sinh(z\sqrt{2\alpha})}\int_0^z y^{n-2}(\alpha y^2 - \frac{n(n-1)}{2})\cosh(y\sqrt{2\alpha})dy. \\
&= \frac{\sqrt{2\alpha}}{\alpha \sinh(z\sqrt{2\alpha})}\Big(\int_0^z \alpha y^n \cosh(y\sqrt{2\alpha})dy \\
&\qquad\qquad - \int_0^z \frac{n(n-1)}{2} y^{n-2}\cosh(y\sqrt{2\alpha})dy\Big) \qquad (28)
\end{aligned}
$$

By partial integration the first integral in (28) equals

$$
\frac{\alpha z^n}{\sqrt{2\alpha}}\sinh(z\sqrt{2\alpha}) - \frac{n\sqrt{2\alpha}}{2}\int_0^z y^{n-1}\sinh(y\sqrt{2\alpha})dy,
$$

and the second one

$$
\frac{nz^{n-1}}{2}\cosh(z\sqrt{2\alpha}) - \frac{n\sqrt{2\alpha}}{2}\int_0^z y^{n-1}\sinh(y\sqrt{2\alpha})dy.
$$

Consequently,

$$
Q_n(z) = z^{n-1}(z - \frac{n}{\sqrt{2\alpha}}\coth(z\sqrt{2\alpha})). \qquad (29)
$$

It is seen that the function $z \mapsto Q_n(z)$ has a unique positive root x^* and is non-decreasing for $z \geq x^*$, where x^* is the solution of the equation

$$
z - \frac{n}{\sqrt{2\alpha}}\coth(z\sqrt{2\alpha}) = 0,
$$

equivalently,

$$
z\sinh(z\sqrt{2\alpha}) - \frac{n}{\sqrt{2\alpha}}\cosh(z\sqrt{2\alpha}) = 0.
$$

Putting $\hat{f}_n = Q_n \vee 0$ it is seen that $\hat{f}_n$ fulfils the conditions of Theorem 2.5 in Article I.

Case 2. $n = 1$. As showed in identity (a) of Proposition 6.2 the function $\tilde{f}_1$ does not satisfy equality (26). We overcome this difficulty by showing that there exists a point y^* such that the function $\tilde{g}_1(x) := \tilde{f}_1(x)1_{\{x\geq y^*\}}$ has the following properties

(i) $\frac{1}{\alpha}\mathbb{E}_x(\tilde{g}_1(X_T)1_{\{X_T\geq y^*\}}) = x \quad \text{for} \quad x \geq y^*,$

(ii) $\frac{1}{\alpha}\mathbb{E}_x(\tilde{g}_1(X_T)1_{\{X_T\geq y^*\}}) \geq x \quad \text{for} \quad 0 < x \leq y^*.$

If the function

$$P_1(z) := \frac{1}{\alpha}\int_{y^*}^{z} \tilde{g}_1(y)\mathbb{P}_x(X_T \in dy|M_T = z)$$

fulfils condition (a) of Theorem 2.5 in Article I then $\hat{f}_1(z) := P_1(z)1_{\{z\geq y^*\}}$ is the needed function. We now prove properties (i) and (ii). Using formula 1.0.4 in [13, p 333])

$$\mathbb{P}_x(X_T \in dz) = \frac{\sqrt{\alpha}}{\sqrt{2}}(e^{-|z-x|\sqrt{2\alpha}} + e^{-(z+x)\sqrt{2\alpha}})$$

we obtain for an arbitrary $x_o > 0$

$$\begin{aligned}\frac{1}{\alpha}\mathbb{E}_x(\tilde{g}_1(X_T)1_{\{X_T\geq x_o\}}) &= \mathbb{E}_x(X_T 1_{\{X_T\geq x_o\}}) \\ &= \frac{\sqrt{\alpha}}{\sqrt{2}}\int_{x_o}^{\infty} z(e^{-|z-x|\sqrt{2\alpha}} + e^{-(z+x)\sqrt{2\alpha}})dz.\end{aligned}$$

If $x \geq x_o$

$$\begin{aligned}\frac{1}{\alpha}\mathbb{E}_x(\tilde{g}_1(X_T)1_{\{X_T\geq x_o\}}) &= \frac{\sqrt{\alpha}}{\sqrt{2}}\int_{x_o}^{x} z(e^{(z-x)\sqrt{2\alpha}} + e^{-(z+x)\sqrt{2\alpha}})dz \\ &\quad + \frac{\sqrt{\alpha}}{\sqrt{2}}\int_{x}^{\infty} z(e^{-(z-x)\sqrt{2\alpha}} + e^{-(z+x)\sqrt{2\alpha}})dz \\ &= \frac{\sqrt{\alpha}}{\sqrt{2}}\Big[2e^{-x\sqrt{2\alpha}}\int_{x_o}^{x} z\cosh(z\sqrt{2\alpha})dz \\ &\qquad + 2\cosh(x\sqrt{2\alpha})\int_{x}^{\infty} ze^{-z\sqrt{2\alpha}}dz\Big]. \end{aligned} \tag{30}$$

Using the integration by parts the first integral in (30) equals

$$\begin{aligned}&\frac{x}{\sqrt{2\alpha}}\sinh(x\sqrt{2\alpha}) - \frac{x_o}{\sqrt{2\alpha}}\sinh(x_o\sqrt{2\alpha}) \\ &\quad - \frac{1}{2\alpha}\cosh(x\sqrt{2\alpha}) + \frac{1}{2\alpha}\cosh(x_o\sqrt{2\alpha}).\end{aligned}$$

Consequently, since

$$\int_{x}^{\infty} ze^{-z\sqrt{2\alpha}}dz = \frac{x}{\sqrt{2\alpha}}e^{-x\sqrt{2\alpha}} + \frac{1}{2\alpha}e^{-x\sqrt{2\alpha}},$$

we have

$$\frac{1}{\alpha}\mathbb{E}_x(\tilde{g}_1(X_T)1_{\{X_T\geq x_o\}}) = \sqrt{\frac{\alpha}{2}}\Big\{2e^{-x\sqrt{2\alpha}}\Big[\Big(\frac{x}{\sqrt{2\alpha}}\sinh(x\sqrt{2\alpha}) - \frac{1}{2\alpha}\cosh(x\sqrt{2\alpha})\Big) + \Big(\frac{1}{2\alpha}\cosh(x_o\sqrt{2\alpha}) - \frac{x_o}{\sqrt{2\alpha}}\sinh(x_o\sqrt{2\alpha})\Big)\Big] + 2\cosh(x\sqrt{2\alpha})\Big(\frac{x}{\sqrt{2\alpha}}e^{-x\sqrt{2\alpha}} + \frac{1}{2\alpha}e^{-x\sqrt{2\alpha}}\Big)\Big\}$$

$$= x + \sqrt{2\alpha}e^{-x\sqrt{2\alpha}}\Big(\frac{1}{2\alpha}\cosh(x_o\sqrt{2\alpha}) - \frac{x_o}{\sqrt{2\alpha}}\sinh(x_o\sqrt{2\alpha})\Big).$$

Choosing now $x_o = y^*$ with y^* as the unique solution of the equation

$$y\sinh(y\sqrt{2\alpha}) - \frac{1}{\sqrt{2\alpha}}\cosh(y\sqrt{2\alpha}) = 0 \tag{31}$$

yields

$$\frac{1}{\alpha}\mathbb{E}_x(\tilde{g}_1(X_T)1_{\{X_T\geq y^*\}}) = x.$$

For $0 < x < y^*$

$$\begin{aligned}\frac{1}{\alpha}\mathbb{E}_x(\tilde{f}(X_T)1_{\{X_T\geq y^*\}}) &= \sqrt{\frac{\alpha}{2}}\int_{y^*}^{\infty} z(e^{-(z-x)\sqrt{2\alpha}} + e^{(z+x)\sqrt{2\alpha}})dz \\ &= \sqrt{\frac{\alpha}{2}}2\cosh(x\sqrt{2\alpha})\int_{y^*}^{\infty} ze^{-z\sqrt{2\alpha}}dz \\ &= \cosh(x\sqrt{2\alpha})e^{-y^*\sqrt{2\alpha}}(y^* + \frac{1}{\sqrt{2\alpha}}) \\ &= \cosh(x\sqrt{2\alpha})\frac{y^*}{\cosh(y^*\sqrt{2\alpha})} \geq x,\end{aligned} \tag{32}$$

where we used the fact that y^* is the solution of (31) and, hence, satisfies

$$e^{-y^*\sqrt{2\alpha}}(y^* + \frac{1}{\sqrt{2\alpha}}) = \frac{y^*}{\cosh(y^*\sqrt{2\alpha})}.$$

Inequality (32) is implied from the fact that the function $x \mapsto \dfrac{x}{\cosh(x\sqrt{2\alpha})}$ is increasing on $(0, y^*)$. Thus, $\tilde{g}_1$ satisfies conditions (i) and (ii) above. Using

(27) we now obtain the needed function $\hat{f}_1$

$$\hat{f}_1(z) = P_1(z)1_{\{z\geq y^*\}} = \begin{cases} \dfrac{\sqrt{2\alpha}}{\alpha\sinh(z\sqrt{2\alpha})}\displaystyle\int_{y^*}^{z} y\ \cosh(y\sqrt{2\alpha})dy, & z \geq y^*, \\ 0, & z < y^*. \end{cases}$$

$$= \begin{cases} z - \frac{1}{\sqrt{2\alpha}}\coth(z\sqrt{2\alpha}), & z \geq y^*, \\ 0, & z < y^*. \end{cases}$$

Notice that the function P_1 is obtained from (29) formally by putting therein $n = 1$.

6.2 Two-sided problem for Brownian motion

Article I contains one example on two-sided problem. In this example the underlying process is a continuous time Markov chain. Here, we give an additional example and show how the method in Article I can be applied to solve OSP (16) for Brownian motion and the reward $G(x) = |x|^n$, $n \geq 1$. Let $\hat{M}_T := \sup_{0\leq t\leq T} |B_t|$ denote the maximum of reflecting Brownian motion up to time T and $\tilde{f}_n$ be as in (25). Consider first the case $n \geq 2$. From (28) and (29) we have

$$\begin{aligned} Q_n(z) &= \frac{1}{\alpha}\int_0^z \tilde{f}_n(y)\mathbb{P}_x(|B_T| \in dy|\hat{M}_T = z) \\ &= z^{n-1}(z - \frac{n}{\sqrt{2\alpha}}\coth(z\sqrt{2\alpha})). \end{aligned}$$

Recall that Q_n has a unique positive root x^*, is increasing and positive on $(x^*, +\infty)$ and satisfies for all x

$$\mathbb{E}_x(Q_n(\hat{M}_T)) = \frac{1}{\alpha}\mathbb{E}_x(\tilde{f}_n(|B|_T)) = x^n.$$

Let

$$\hat{Q}_n(z) := \begin{cases} Q_n(z) & \text{if } z \geq 0, \\ Q_n(-z) & \text{if } z \leq 0. \end{cases}$$

Clearly, the function $\hat{Q}_n$ has two non-zero roots $-x^*$ and x^*. Furthermore, $\hat{Q}_n$ is increasing on $(x^*, +\infty)$ and decreasing on $(-\infty, -x^*)$. We claim that

$$Q_n(a)1_{\{a\geq x^*\}} \vee Q_n(-b)1_{\{b\leq -x^*\}} = Q_n(a \vee -b)1_{\{a\vee -b\geq x^*\}}. \tag{33}$$

Indeed,

$$\begin{aligned} Q_n(a)1_{\{a\geq x^*\}} \vee Q_n(-b)1_{\{b\leq -x^*\}} &= Q_n(a)1_{\{a\geq x^*\}}1_{\{a\geq -b\}} + Q_n(-b)1_{\{b\leq -x^*\}}1_{\{a\leq -b\}} \\ &= Q_n(a\vee -b)1_{\{a\vee -b\geq x^*\}}. \end{aligned}$$

From the spatial symmetry of Brownian motion it follows that

$$M_T \vee -I_T = \hat{M}_T.$$

where $M_T := \sup_{0\leq t\leq T} B_t$ and $I_T := \inf_{0\leq t\leq T} B_t$. Using this and (33) we get

$$\begin{aligned} &\mathbb{E}_x\Big(\hat{Q}_n(M_T)1_{\{M_T\geq x^*\}} \vee \hat{Q}_n(I_T)1_{\{I_T\leq -x^*\}}\Big) \\ &\qquad = \mathbb{E}_x\Big(Q_n(M_T)1_{\{M_T\geq x^*\}} \vee Q_n(I_T)1_{\{I_T\leq -x^*\}}\Big) \\ &\qquad = \mathbb{E}_x\Big(Q_n(M_T\vee -I_T)1_{\{M_T\vee -I_T\geq x^*\}}\Big) \\ &\qquad = \mathbb{E}_{|x|}\Big(Q_n(\hat{M}_T)1_{\{\hat{M}_T\geq x^*\}}\Big) \\ &\qquad \begin{cases} = |x|^n, & |x|\geq x^*, \\ \geq |x|^n, & |x|\leq x^*. \end{cases} \end{aligned}$$

Let $\hat{f}_n = \hat{Q}_n \vee 0$. Then it holds

$$\begin{aligned} \mathbb{E}_x\Big(\sup_{0\leq t\leq T} \hat{f}_n(B_t)\Big) = &\mathbb{E}_x\Big(Q_n(M_T\vee -I_T)1_{\{M_T\vee -I_T\geq x^*\}}\Big) \\ &\begin{cases} = |x|^n, & |x|\geq x^*, \\ \geq |x|^n, & |x|\leq x^*. \end{cases} \end{aligned}$$

i.e., condition (b) of Theorem 2.7 in Article I is valid. Clearly also condition (a) holds and, hence, the solution of OSP (16) is given by

$$V(x) = \mathbb{E}_x\Big(\hat{f}(M_T)1_{\{M_T\geq x^*\}} \vee \hat{f}(I_T)1_{\{I_T\leq -x^*\}}\Big).$$

For the case $n = 1$, choosing the functions $Q_1(z) \equiv P_1(z)1_{\{z\geq y^*\}}$ as in section 6.2, i.e.,

$$Q_1(z) = \begin{cases} z - \frac{1}{\sqrt{2\alpha}}\coth(z\sqrt{2\alpha}), & z\geq y^*, \\ 0, & z < y^*, \end{cases}$$

where y^* is the unique solution of (31), we have similarly

$$\hat{f}_1(z) = \hat{Q}_1(z)(1_{\{z \le -y^*\}} + 1_{\{z \ge y^*\}}),$$

where

$$\hat{Q}_1(z) := \begin{cases} Q_1(z) & \text{if } z \ge 0, \\ Q_1(-z) & \text{if } z \le 0. \end{cases}$$

7 Summaries of the included articles I-IV

7.1 Article I: Optimal stopping of strong Markov processes

In this paper we present a new technique to solve OSP which applies to general Hunt processes, in particular, both for diffusions and Lévy processes. The motivation is inspired by recent studies for Lévy processes in which the methodology to derive the solutions is essentially based on the Wiener-Hopf factorization. The crucial ingredient in our approach is to use the following representation of α-excessive functions as expected suprema (see [25], for related works, see also [7, 8, 23, 24]):

Let X be a Hunt process on the state space $\mathbb{R}$ and $f : \mathbb{R} \mapsto \mathbb{R}_+$ an upper semicontinuous function and define

$$u(x) := \mathbb{E}_x\Big(\sup_{0 \le t \le T} f(X_t)\Big).$$

Then the function $u : \mathbb{R} \mapsto \mathbb{R}_+ \cup \{+\infty\}$ is α-excessive.

Furthermore, it is seen that the Wiener-Hopf factorization used for Lévy processes is in our general framework replaced by a path decomposition of the underlying strong Markov process.

Let T be an exponential distributed random variable with parameter $\alpha > 0$, independent of X. Consider OSP

$$V(x) = \sup_{\tau \in \mathcal{M}} \mathbb{E}_x(e^{-\alpha\tau} G(X_\tau)) = \mathbb{E}_x(e^{-\alpha\tau^*} G(X_{\tau^*})),$$

The main result is as follows: assume that there exists an upper semicontinuous function $\hat{f} : S \mapsto \mathbb{R}$ and a point $x^* \in S$ such that

(a) (i) $\hat{f}(x) \le 0$ for $x \le x^*$,

(ii) $\hat{f}(x)$ is positive and non-decreasing for $x < x^*$.

(b) (i) $\mathbb{E}_x(\sup_{0\le t\le T}\hat{f}(X_t)) = G(x)$ for $x \ge x^*$,

(ii) $\mathbb{E}_x(\sup_{0\le t\le T}\hat{f}(X_t)) \ge G(x)$ for $x < x^*$.

Then the value function V of the optimal stopping problem is given by

$$V(x) = \mathbb{E}_x\Big(\sup_{0\le t\le T}\hat{f}(X_t)1_{\{X_t\ge x^*\}}\Big) = \mathbb{E}_x\Big(\hat{f}(M_T)1_{\{M_T\ge x^*\}}\Big)$$

and

$$\tau^* := \inf\{t \ge 0 : X_t \ge x^*\}$$

is an optimal stopping time.

Moreover, we discuss how to find the function $\hat{f}$ and present some examples. We are also able to connect this new approach with the approach based on the Riesz representation of excessive functions. We refer also to Section 6 above for two more examples of the approach.

7.2 Article II: Differentiability of excessive functions of one-dimensional diffusions and the principle of smooth fit

In this article we are interested in studying the differentiability of excessive functions of one-dimensional diffusions and in particular, the principle of smooth fit for diffusions with non-smooth characteristics. We obtain the following main results:

Result I: Let u be an excessive function for a diffusion process X on the state space I with the smooth characteristics and F be an increasing continuous function on I. Assume that for $\alpha > 0$, the fundamental solutions ψ_α and φ_α are F-differentiable at a point $z \in I$. Then

$$\frac{d^-u}{dF}(z) - \frac{d^+u}{dF}(z) \ge 0.$$

Moreover, u is F-differentiable at z if and only if $\sigma_u(\{z\}) = 0$, where σ_u is the representing measure in the Riesz representation of u.

As an application of this result to optimal stopping problems we obtain the following extended version of the principle of smooth fit: if the functions G, φ_α and ψ_α are F-differentiable at the stopping point x^* then the value function V is also F-differentiable at x^* and

$$\frac{dV}{dF}(x^*) = \frac{dG}{dF}(x^*).$$

In particular, choosing $F \equiv S$, the scale function of X, we obtain the condition of the scale smooth fit as derived in [48]. When $F(x) \equiv x$ we obtain the condition for the regular smooth fit which coincides with the result in [41].

Result II: We extend the result I for diffusions with sticky points, i.e., the speed measure has atoms. As a specific example, we study the principle of smooth fit for optimal stopping of sticky Brownian motion. The size of the jump in the derivative of an α-excessive function at the sticky point can be decomposed into two factors. More precisely, let $X = (X_t)_{t\geq 0}$ be a one-dimensional regular diffusion process and u be an α-excessive function of X. Then

$$u^-(z) - u^+(z) = \sigma_u(\{z\}) - m(\{z\})\alpha u(z), \tag{34}$$

where $u^-(u^+)$ denotes the left (right) derivative with respect to the scale function S. This fact is employed to study the optimal stopping problem of Brownian motion sticky at 0 and the reward function $G(x) = (x+1)^+$. It is seen that, cf. [18], if the discounting parameter $\alpha \in [\alpha_1, \alpha_2]$ then the sticky point 0 is a boundary point of the stopping region. Furthermore, when $\alpha = \alpha_2$ the smooth fit holds at 0, otherwise, it fails. Using formula (34), we decompose the value of the size of the jump in the derivative of the value function into the components described in (34).

7.3 Article III: Probabilistic approach to Appell polynomials

This article is about the probabilistic approach to Appell polynomials. In particular, we study in this framework some classical Appell polynomials: the Bernoulli, Euler, Hermite and Laguerre polynomials. An important property deduced from the probabilistic approach is the so-called mean value property from which many other properties can be derived. We find some conditions for the underlying random variable which guarantee that the Appell polynomial associated with this variable has the moment representation. We also show that the probabilistic approach provides us with powerful tools to reprove some old results, e.g., that the Bernoulli numbers and the Euler numbers have alternating signs.

Since the existence and uniqueness of positive root of Appell polynomials play a crucial role in studying optimal stopping of Lévy processes, we also give sufficient conditions for this property to hold. An interesting observation is that for the gamma distributed random variable $\xi \sim \Gamma(a, 1)$ the Appell polynomials associated with ξ have a unique positive root if and only if $0 < a \leq 1$.

7.4 Article IV: A note on the generalized Bernoulli and Euler polynomials

In this article we extend the main results due to Srivastava and Pintér [55] using the probabilistic approach to Appell polynomials. In particular, the mean value property is applied. We formulate a relationship between the generalized Bernoulli polynomials and the generalized Euler polynomials. As a further application of this approach we also give new proofs for the results obtained in [14] and [55].

Bibliography

[1] L. Alili and A. E. Kyprianou. Some remarks on first passage of Lévy processes, the American put and pasting principles. *Ann. Appl. Probab.*, 15(3):2062–2080, 2005.

[2] L. H. R. Alvarez. A class of solvable singular stochastic control problems. *Stochastics Stochastics Rep.*, 67(1-2):83–122, 1999.

[3] L. H. R. Alvarez. On the properties of r-excessive mappings for a class of diffusions. *Ann. Appl. Probab.*, 13(4):1517–1533, 2003.

[4] L. H. R. Alvarez and P. Salminen. Optimal stopping of linear diffusions: A synthesis. *Under preperation.*

[5] P. Appell. Sur une classe de polynômes. *Ann. Sci. École Norm. Sup. (2)*, 9:119–144, 1880.

[6] D. Applebaum. *Lévy processes and stochastic calculus.* Cambridge University Press, Cambridge, 2004.

[7] P. Bank and N. El Karoui. A stochastic representation theorem with applications to optimization and obstacle problems. *Ann. Probab.*, 32(1B):1030–1067, 2004.

[8] P. Bank and H. Föllmer. American options, multi-armed bandits, and optimal consumption plans: a unifying view. In *Paris-Princeton Lectures on Mathematical Finance, 2002*, volume 1814 of *Lecture Notes in Math.*, pages 1–42. Springer, Berlin, 2003.

[9] R. F. Bass. *Stochastic processes.* Cambridge University Press, Cambridge, 2011.

[10] J. Bertoin. *Lévy processes.* Cambridge University Press, Cambridge, 1996.

[11] R. M. Blumenthal and R. K. Getoor. *Markov processes and potential theory.* Academic Press, New York, 1968.

[12] R. P. Boas, Jr. The Stieltjes moment problem for functions of bounded variation. *Bull. Amer. Math. Soc.*, 45(6):399–404, 1939.

[13] A. N. Borodin and P. Salminen. *Handbook of Brownian motion—facts and formulae.* Birkhäuser Verlag, Basel, second edition, 2002.

[14] Gi-S. Cheon. A note on the Bernoulli and Euler polynomials. *Appl. Math. Lett.*, 16(3):365–368, 2003.

[15] S. Christensen and A. Irle. A note on pasting conditions for the American perpetual optimal stopping problem. *Statist. Probab. Lett.*, 79(3):349–353, 2009.

[16] S. Christensen and P. Salminen. Riesz representation and optimal stopping with two case studies. *Preprint arXiv:1309.2469.*

[17] K. L. Chung and J. B. Walsh. *Markov processes, Brownian motion, and time symmetry.* Springer, New York, second edition, 2005.

[18] F. Crocce and E. Mordecki. Explicit solutions of one-sided optimal stopping problems for one-dimensional diffusions. *to appear in Stochastics, DOI:10.1080/17442508.2013.837467*, 2013.

[19] S. Dayanik and I. Karatzas. On the optimal stopping problem for one-dimensional diffusions. *Stochastic Processes Appl.*, 107:173–212, 2003.

[20] E. B. Dynkin. Optimal choice of the stopping moment of a Markov process. *Dokl. Akad. Nauk SSSR*, 150:238–240, 1963.

[21] E. B. Dynkin. *Markov processes. Vols. II*, volume 122. Academic Press Inc., Publishers, New York, 1965.

[22] E. B. Dynkin and A. A. Yushkevich. *Markov processes: Theorems and problems.* Plenum Press, New York, 1969.

[23] N. El Karoui and H. Föllmer. A non-linear Riesz representation in probabilistic potential theory. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(3):269–283, 2005.

[24] N. El Karoui and A. Meziou. Max-Plus decomposition of supermartingales and convex order. Application to American options and portfolio insurance. *Ann. Probab.*, 36(2):647–697, 2008.

[25] H. Föllmer and T. Knispel. A representation of excessive functions as expected suprema. *Probab. Math. Statist.*, 26(2):379–394, 2006.

[26] D. Freedman. *Brownian motion and diffusion.* Holden-Day, San Francisco, Calif., 1971.

[27] B. I. Grigelionis and A. N. Shiryaev. On the Stefan problem and optimal stopping rules for Markov processes. *Teor. Verojatnost. i Primenen*, 11:612–631, 1966.

[28] K. Itô and H. P. McKean, Jr. *Diffusion processes and their sample paths.* Springer-Verlag, Berlin, 1974.

[29] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical methods for financial markets.* Springer Finance. Springer-Verlag London Ltd., London, 2009.

[30] I. Karatzas and S. E. Shreve. Connections between optimal stopping and singular stochastic control. I. Monotone follower problems. *SIAM J. Control Optim.*, 22(6):856–877, 1984.

[31] I. Karatzas and S. E. Shreve. Connections between optimal stopping and singular stochastic control. II. Reflected follower problems. *SIAM J. Control Optim.*, 23(3):433–451, 1985.

[32] S. Karlin and H. M. Taylor. *A second course in stochastic processes.* Academic Press Inc., New York, 1981.

[33] H. Kunita and T. Watanabe. Markov processes and Martin boundaries. I. *Illinois J. Math.*, 9:485–526, 1965.

[34] A. E. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications.* Springer-Verlag, Berlin, 2006.

[35] A. E. Kyprianou and B. A. Surya. On the Novikov-Shiryaev optimal stopping problems in continuous time. *Electron. Comm. Probab.*, 10:146–154 (electronic), 2005.

[36] P. Mandl. *Analytical treatment of one-dimensional Markov processes.* Springer-Verlag, 1968.

[37] R. C. Merton. Theory of rational option pricing. *Bell J. Econom. and Management Sci.*, 4:141–183, 1973.

[38] E. Mordecki and P. Salminen. Optimal stopping of Hunt and Lévy processes. *Stochastics*, 79(3-4):233–251, 2007.

[39] A. Novikov and A. Shiryaev. On an effective case of the solution of the optimal stopping problem for random walks. *Teor. Veroyatn. Primen.*, 49(2):373–382, 2004.

[40] B. Øksendal. *Stochastic differential equations.* Springer-Verlag, Berlin, sixth edition, 2003.

[41] G. Peskir. Principle of smooth fit and diffusions with angles. *Stochastics*, 79(3-4):293–302, 2007.

[42] G. Peskir and A. Shiryaev. *Optimal stopping and free-boundary problems.* Lectures in Mathematics, ETH Zürich. Basel: Birkhäuser, 2006.

[43] P. Picard and C. Lefèvre. The probability of ruin in finite time with discrete claim size distribution. *Scand. Actuar. J.*, (1):58–69, 1997.

[44] P. Picard, C. Lefèvre, and I. Coulibaly. Problèmes de ruine en théorie du risque à temps discret avec horizon fini. *J. Appl. Probab.*, 40(3):527–542, 2003.

[45] E. D. Rainville. *Special functions.* Chelsea Publishing Co., Bronx, N.Y., first edition, 1971.

[46] D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* Springer-Verlag, Berlin, third edition, 1999.

[47] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2.* John Wiley & Sons Inc., New York, 1987.

[48] P. Salminen. Optimal stopping of one-dimentional diffusions. *Math. Nachr*, 124:85–101, 1985.

[49] K. Sato. *Lévy processes and infinitely divisible distributions.* Cambridge University Press, Cambridge, 1999.

[50] W. Schoutens. *Stochastic processes and orthogonal polynomials*, volume 146 of *Lecture Notes in Statistics.* Springer-Verlag, New York, 2000.

[51] I. M. Sheffer. Some properties of polynomial sets of type zero. *Duke Math. J.*, 5:590–622, 1939.

[52] I. M. Sheffer. Note on Appell polynomials. *Bull. Amer. Math. Soc.*, 51:739–744, 1945.

[53] A. N. Shiryaev. *Optimal stopping rules.* Springer-Verlag, Berlin, 2008.

[54] J. L. Solé and F. Utzet. Time-space harmonic polynomials relative to a Lévy process. *Bernoulli*, 14(1):1–13, 2008.

[55] H. M. Srivastava and Á. Pintér. Remarks on some relationships between the Bernoulli and Euler polynomials. *Appl. Math. Lett.*, 17(4):375–380, 2004.

[56] H. M. Taylor. Optimal stopping in a Markov process. *Ann. Math. Statist.*, 39:1333–1344, 1968.

[57] A. Wald and J. Wolfowitz. Optimum character of the sequential probability ratio test. *Ann. Math. Statistics*, 19:326–339, 1948.

[58] D. Widder. *The Laplace Transform.* Princeton University Press, 1941.

ISBN 978-952-12-3045-5