Lecture Notes Series No. 68, 2001, *Dept. Math. Univ. Aarhus*

Principles of Optimal Stopping and Free-Boundary Problems

GORAN PESKIR

Department of Mathematical Sciences
University of Aarhus
Denmark

Denne afhandling er, i forbindelse med de nedenfor anførte, tidligere offentliggjorte afhandlinger, af Det naturvidenskabelige Fakultet ved Aarhus Universitet antaget til forsvar for den naturvidenskabelige doktorgrad.

Forsvarshandlingen finder sted fredag den 12. april 2002 kl. 13.15 i Auditorium F på Institut for Matematiske Fag, Aarhus Universitet.

Aarhus, den 9. januar 2002

Karl Pedersen

Dekan

CONTENTS

PREFACE

This thesis consists of twenty-two papers that I wrote on the subject in the period 1994-2000. A short description of each paper with some historical remarks is given in Chapter 1. The four most illustrative papers are presented in Chapters 2-5. These chapters also contain 17 figures.

My attraction to optimal stopping problems, to which I was introduced by Albert N. Shiryaev, was twofold. Firstly, a search for the solution forced me to consider motion along a sample path, which consequently increased my understanding of stochastic processes. Secondly, I admired the connection to elliptic and parabolic equations of mathematical physics, and striving for boundary conditions when tackling a problem was always exciting. I still regard this connection as one of the most fascinating accomplishments of modern mathematics and theory of probability.

My advantage in dealing with optimal stopping problems was, as I see it now, that I was not fully aware of the general theory of optimal stopping. The only thing I knew, or was dealing with, was the many examples. It then took me a while to realise, set aside technicalities, that the basic principles are few. During this process I had the pleasure of collaborating with Albert N. Shiryaev, Svend Erik Graversen, Jesper L. Pedersen and Jørgen Hoffmann-Jørgensen, and I thank them all for the many discussions we had during those years.

Thanks are also due to my colleagues at the Institute of Mathematics, University of Aarhus for the long-term support. Financial support from the Centre for Mathematical Physics and Stochastics (Denmark) and the Centre for Analytical Finance (Aarhus) is gratefully acknowledged.

Aarhus, January 2001 Goran Peskir

LIST OF PAPERS

[P1] Optimal stopping of the maximum process: The maximality principle. *Research Report* No. 377, 1997, *Dept. Theoret. Statist. Aarhus* (30 pp). *Ann. Probab.* Vol. 26, No. 4, 1998, (1614-1640).

[P2] Stopping Brownian motion without anticipation as close as possible to its ultimate maximum. *Research Report* No. 405, 1999, *Dept. Theoret. Statist. Aarhus*, (with S. E. Graversen and A. N. Shiryaev), (11 pp). *Theory Probab. Appl.* Vol. 45, No. 1, 2000, (125-136).

[P3] Sequential testing problems for Poisson processes. *Research Report* No. 400, 1998, *Dept. Theoret. Statist. Aarhus*, (with A. N. Shiryaev), (20 pp). *Ann. Statist.* Vol. 28, No. 3, 2000, (837-859).

[P4] Solving the Poisson disorder problem. *Research Report* No. 419, 2000, *Dept. Theoret. Statist. Aarhus*, (with A. N. Shiryaev), (17 pp). *Advances in Finance and Stochastics: Essays in Honour of Dieter Sondermann*, Springer-Verlag, 2002, (295-312).

[P5] On Doob's maximal inequality for Brownian motion. *Research Report* No. 337 (1995), *Dept. Theoret. Statist. Aarhus*, (with S. E. Graversen), (13 pp). *Stochastic Process. Appl.* Vol. 69, No. 1, 1997, (111-125).

[P6] The best Doob-type bounds for the maximum of Brownian paths. *Research Report* No. 363, 1996, *Dept. Theoret. Statist. Aarhus* (9 pp). *Proc. High Dimensional Probability I* (Oberwolfach 1996), *Progr. Probab.* Vol. 43, Birkhäuser Basel, 1998, (287-296).

[P7] Optimal stopping in the $L \log L$-inequality of Hardy and Littlewood. *Research Report* No. 360, 1996, *Dept. Theoret. Statist. Aarhus*, (with S. E. Graversen), (12 pp). *Bull. London Math. Soc.* Vol. 30, No. 2, 1998, (171-181).

[P8] The integral analogue of the Hardy-Littlewood $L \log L$-inequality for Brownian motion. *Research Report* No. 359, 1996, *Dept. Theoret. Statist. Aarhus* (12 pp). *Math. Inequal. Appl.* Vol. 1, No. 1, 1998, (137-148).

[P9] Designing options given the risk: The optimal Skorokhod-embedding problem. *Research Report* No. 389, 1997, *Dept. Theoret. Statist. Aarhus* (18 pp). *Stochastic Process. Appl.* Vol. 81, No. 1, 1999, (25-38).

[P10] On the Russian option: The expected waiting time. *Research Report* No. 325, 1995, *Dept. Theoret. Statist. Aarhus*, (with S. E. Graversen), (11 pp). *Theory Probab. Appl.* Vol. 42, No. 3, 1997, (564-575).

[P11] Optimal stopping and maximal inequalities for geometric Brownian motion. *Research Report* No. 330 (1995), *Dept. Theoret. Statist. Aarhus*, (with S. E. Graversen), (18 pp). *J. Appl. Probab.* Vol. 35, No. 4, 1998, (856-872).

[P12] Optimal stopping and maximal inequalities for linear diffusions. *Research Report* No. 335 (1995), *Dept. Theoret. Statist. Aarhus*, (with S. E. Graversen), (18 pp). *J. Theoret. Probab.* Vol. 11, No. 1, 1998, (259-277).

[P13] Solving non-linear optimal stopping problems by the method of time-change. *Research Report* No. 390, 1997, *Dept. Theoret. Statist. Aarhus*, (with J. L. Pedersen), (24 pp). *Stochastic Anal. Appl.* Vol. 18, No. 5, 2000, (811-835).

[P14] Optimal stopping inequalities for the integral of Brownian paths. *Research Report* No. 355 (1996), *Dept. Theoret. Statist. Aarhus* (9 pp). *J. Math. Anal. Appl.* Vol. 222, No. 1, 1998, (244-254).

[P15] On Wald-type optimal stopping for Brownian motion. *Math. Inst. Aarhus, Preprint Ser.* No. 10, 1994, (with S. E. Graversen), (15 pp). *J. Appl. Probab.* Vol. 34, No. 1, 1997, (66-73).

[P16] On the Brownian first-passage time over a one-sided stochastic boundary. *Research Report* No. 371, 1997, *Dept. Theoret. Statist. Aarhus*, (with A. N. Shiryaev), (11 pp). *Theory Probab. Appl.* Vol. 42, No. 3, 1997, (591-602).

[P17] Computing the expectation of the Azéma-Yor stopping time. *Research Report* No. 364, 1996, *Dept. Theoret. Statist. Aarhus*, (with J. L. Pedersen), (10 pp). *Ann. Inst. H. Poincaré Probab. Statist.* Vol. 34, No. 2, 1998, (265-276).

[P18] The Azéma-Yor embedding in Brownian motion with drift. *Research Report* No. 398, 1998, *Dept. Theoret. Statist. Aarhus* (12 pp). *Proc. High Dimensional Probability II* (Seattle 1999), *Progr. Probab.* Vol. 47, Birkhäuser Boston, 2000, (207-221).

[P19] Bounding the maximal height of a diffusion by the time elapsed. *Research Report* No. 412, 1999, *Dept. Theoret. Statist. Aarhus* (9 pp). *J. Theoret. Probab.* Vol. 14, No. 3, 2001, (845-855).

[P20] Maximal inequalities for the Ornstein-Uhlenbeck process. *Research Report* No. 393, 1998, *Dept. Theoret. Statist. Aarhus*, (with S. E. Graversen), (8 pp). *Proc. Amer. Math. Soc.* Vol. 128, No. 10, 2000, (3035-3041).

[P21] Controlling the velocity of Brownian motion by its terminal value. *Research Report* No. 391, 1998, *Dept. Theoret. Statist. Aarhus* (11 pp). *Analytic Geom. Inequal. Applications, Math. Appl.* Vol. 478, Kluwer Academic Publishers, 1999, (323-333).

[P22] Maximal inequalities for reflected Brownain motion with drift. *Research Report* No. 409, 1999, *Dept. Theoret. Statist. Aarhus*, (with A. N. Shiryaev), (8 pp). To appear in *Theory Probab. Math. Statist.*

1. Introduction

This thesis studies optimal stopping problems for Markov processes by reducing them to free-boundary problems. The aim of this introduction is to say a few words about the origin and history of these and related problems, indicate their interplay and applications, and to present a survey of the papers (Section 1.2) which form the thesis.

For the elegance of the exposition in Section 1.1 we fail to state various regularity conditions in detail. These conditions, if not self-evident, can be found in the references quoted throughout. For the same reason we do not aim at utmost generality but instead concentrate on a few specific examples which are sufficient to illustrate the main ideas.

1.1 General facts

In calculus one is often faced with the problem of finding a minimum of the function $f : I\!R \to I\!R$. The most common way to treat this problem is to solve the equation:

$$f'(x) = 0 \tag{1.1.1}$$

and obtain a candidate x for the point of minimum. Although necessary the condition (1.1.1) is not sufficient. To establish that x is minimising one usually verifies that $f''(x) > 0$.

It should be noted that already this simple example clearly demonstrates that there is nothing mysterious about guessing a candidate for the solution of an optimisation problem. As the next step we also see that a verification procedure must be applied to confirm that the guess is correct.

It turns out that this simple methodology extends to far more general optimization problems. The oldest examples of that type occur in calculus of variations.

Calculus of variations. The original motivations for calculus of variations came from classical physics and geometry (see [11, 12, 27]). The simplest problem in calculus of variations is to solve:

$$\inf_x \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))\, dt \tag{1.1.2}$$

under $x(t_0) = x_0$ and $x(t_1) = x_1$. A necessary condition for $x = x(t)$ to be minimising is that the first variation of the integral over all admissible functions $y = y(t)$ satisfying $y(t_0) = y(t_1) = 0$ be zero, i.e. denoting the integral by $J(x)$ we must have:

$$\delta J(x; y) = 0 \tag{1.1.3}$$

for all such y . This leads to *the Euler equation*:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0 \tag{1.1.4}$$

which under $L_{\dot{x}\dot{x}} \neq 0$ reads as $\ddot{x}(t) = F(t, x(t), \dot{x}(t))$ for some F expressed in terms of L . Recalling that $x(t_0) = x_0$ and $x(t_1) = x_1$ we see that a boundary-value problem for the unknown minimising $x_* = x_*(t)$ is obtained. Functions $x = x(t)$ solving this problem are called extremals.

To rule out some of extremals which are not minimising, one may exploit the fact that the second variation of the integral must be non-negative, i.e. we must have $\delta^2 J(x; y) \geq 0$ for all admissible

y . This leads to the *Jacobi* necessary condition (see [27; p.13]). Other necessary conditions also exist (e.g. the *Legendre* condition [27; p.18] or the *Weierstrass* condition [27; p.35]). A verification procedure is finally applied to show that some of the remaining extremals is minimising.

Over the years the inspiration derived from the classic applications of calculus of variations become greatly diffused. Since 1950 new applications emerged in aerospace sciences, industrial process control, and mathematical economics. This eventually led to the development of the new field of optimal control.

Optimal control. Given the equation of motion:

$$\dot{x}(t) = \mu(t, x(t), u(t)) \tag{1.1.5}$$

with $x(t_0) = x_0$ and $x(t_1) = x_1$, the optimal control problem is to solve:

$$\inf_u \int_{t_0}^{t_1} L(t, x(t), u(t))\, dt \tag{1.1.6}$$

where $u = u(t)$ is a control. This is a *Lagrange* formulation of the problem. There also exist *Mayer* and *Bolza* formulations (just like in classical calculus of variations) that are recalled following (1.1.8) below. The three formulations are known to be equivalent (see e.g. [27; pp.25-26]).

The best known set of necessary conditions for optimality is called *Pontryagin's principle*. It has been formulated during the 1950's (see [13] and [56]). A detailed discussion of this approach is given in [27; Chapter II]. As in classical calculus of variations the optimality question is studied by means of differential properties of mappings into the space of controls.

Bellman [6, 7] initiated a different approach in the 1950's termed *dynamic programming* where the control problem is studied as a function of the initial point. This function is called the *value function*. Given the equation of motion:

$$dX_t = \mu(t, X_t, u_t)\, dt + \sigma(t, X_t, u_t)\, dB_t \tag{1.1.7}$$

with $X_0 = x$, and setting $Z_s = (t+s, X_s)$, the optimal control problem is to solve:

$$V(t, x) = \inf_u E_{t,x}\Big(\int_0^{\tau_D} L(Z_s, u_s)\, ds + M(Z_{\tau_D}) \Big) \tag{1.1.8}$$

where $u = u_t$ is a control, and $\tau_D = \inf\{\, s > 0 \mid Z_s \notin C \,\}$ for some open set $C = D^c$. This is a *Bolza* formulation of the problem (a *Mayer* formulation is obtained when L is identically zero). The process $(B_t)_{t\geq 0}$ appearing in (1.1.7) is a standard Wiener process. It is assumed in (1.1.8) that the Markov process $(Z_s)_{s\geq 0}$ starts at (t, x) under $P_{t,x}$.

A necessary and sufficient condition for optimality (up to a smoothness of V) is that:

$$\frac{\partial V}{\partial t} + \inf_u \Big(\mathbb{L}_X^u(V) + L^u \Big) = 0 \quad \text{in } C \tag{1.1.9}$$

$$V = M \quad \text{on } \partial C \tag{1.1.10}$$

where $\partial/\partial t + \mathbb{L}_X^u$ is the infinitesimal operator of the process $(Z_s)_{s\geq 0}$ with u fixed. The equation (1.1.9) is called the *Hamilton-Jacobi-Bellman equation* (or the *PDE of dynamic programming*).

There is a transparent connection between dynamic programming and Pontryagin's principle (related to a method of solving first-order PDE's) (see [27; Chapter IV, §8]. A corresponding approach to dynamic programming concerned with the simplest problem of calculus of variations is called the *Hamilton-Jacobi* theory (see [16, 17]). Functions appearing in (1.1.8+1.1.9) are related to the *Lagrangian/Hamiltonian* occurring in classical mechanics (see e.g. [35]).

To distinguish a different nature of the problems (1.1.6) and (1.1.8) and equations (1.1.5) and (1.1.7), we speak of deterministic and stochastic optimal control, respectively. Stochastic versions of the principles above (such as stochastic Pontryagin's principle) are still examined (see [74]). Books [27, 28, 74] contain many new and old references on the field of optimal control.

Optimal stopping. A predecessor of dynamic programming is found in Wald's sequential analysis [70] which originated in 1943 (see pp.1-4 in the book for a historical account). Snell [68] formulated a general optimal stopping problem for discrete-time stochastic processes, and using the methods suggested by work of Wald and Wolfowitz [73] and Arrow, Blackwell and Girshick [1], he characterized the value function as the smallest supermartingale dominating the reward sequence (*Snell's envelope*). Studies in this direction ("martingale methods") are summarised in [19].

The key equation $V(x) = \max\{M(x), E_x(V(X_1))\}$ was first stated explicitly in [1; p.214] but was already characterized implicitly by Wald [70]. It is a simplest equation of dynamic programming [6]. This equation is often referred to as *the Wald-Bellman equation* and can be deduced by a dynamic programming principle of "backward induction" (see [65; pp.34-35]). A study of sequential testing problems for continuous-time processes (including Wiener and Poisson process) was initiated by Dvoretzky, Kiefer and Wolfowitz [25], however, with no advance to optimal stopping theory. For more historical details on the discrete-time case see [65; pp.111-112].

Dynkin [26] established the key principle of the optimal stopping theory for Markov processes called the *superharmonic characterization of the payoff*. Given the equation of motion:

$$dX_t = \mu(X_t)\,dt + \sigma(X_t)\,dB_t \tag{1.1.11}$$

the optimal stopping problem is to solve:

$$V(x) = \sup_{\tau} E_x\big(M(X_\tau)\big) \tag{1.1.12}$$

where τ is a stopping time of the process $X = (X_t)_{t\geq 0}$. As above V is called the *value function*, and M is called the *reward* or *gain function*. Dynkin's result then states:

(1.1.13) V is the smallest superharmonic function (relative to X) which dominates M

(1.1.14) $\tau_D = \inf\{\, t > 0 \mid X_t \notin C\,\}$ is the optimal stopping time (the smallest possible)

where $C = \{\, x \mid V(x) > M(x)\,\}$ is the *continuation region*, and $D = C^c$ is the *stopping region*. The set ∂C defines the *optimal stopping boundary*. A function V is *superharmonic* relative to X if $(V(X_t))_{t\geq 0}$ is a supermartingale. It is equivalent to the fact that $\mathbb{L}_X(V) \leq 0$ where $\mathbb{L}_X$ is the infinitesimal operator of X given by:

$$\mathbb{L}_X = \mu\,\frac{\partial}{\partial x} + \frac{\sigma^2}{2}\,\frac{\partial^2}{\partial x^2}\ . \tag{1.1.15}$$

Moreover, the process $(V(X_t))_{0 \le t \le \tau_D}$ is a martingale, and thus $\mathbb{L}_X(V) = 0$ in C .

These facts extend to general Markov processes (taking values in $\mathbb{R}^n$). In such a case $\mathbb{L}_X$ is to be replaced by a more general elliptic or parabolic differential operator, or a second-order integro-differential operator (difference-differential operator) if the process has jumps.

In terms of optimal control theory the problem (1.1.12) is Mayer formulated. A Bolza formulation (time-dependent) is given as follows:

$$V(t,x) = \sup_{\tau} E_{t,x}\bigg(\int_0^{\tau} L(Z_s)\, ds + M(Z_{\tau}) \bigg) \ . \tag{1.1.16}$$

where we set $Z_s = (t{+}s, X_s)$. The infinitesimal operator $\partial/\partial t + \mathbb{L}_X$ of the process $(Z_s)_{s \ge 0}$ is of parabolic type. All said above extends to this case as well.

The fact that the smallest superharmonic majorant V of the gain function M satisfies a Wald-Bellman equation was proved by Grigelionis and Shiryaev [37]. In the continuous-time case one needs to modify the equation by considering the exit times from a small ball around the point. From this equation, going back to (1.1.12), one obtains a *free-boundary problem*:

$$\mathbb{L}_X(V) = 0 \quad \text{in} \ \ C \tag{1.1.17}$$

$$V = M \quad \text{on} \ \ \partial C \tag{1.1.18}$$

where along the value function V , the set C i.e. its boundary ∂C is unknown.

A quick way to derive (1.1.17+1.1.18) is to exploit the well-known connection between the Wiener process and Dirichlet problems which goes back at least to Kakutani [42]. If the supremum in (1.1.12) is attained at the exit time from an open set, then (1.1.17) follows clearly by applying the (strong) Markov property. This reasoning extends to Poisson problems (non-zero right-hand side in (1.1.17)) and to a large number of additive (integral) and multiplicative (exponential) functionals yielding all elliptic and parabolic equations (of second-order). On the other hand, a basic fact of stochastic calculus (Itô formula) leads to the following key identity:

$$V(X_t) = V(X_0) + \int_0^t \mathbb{L}_X(X_s)\, ds + M_t \tag{1.1.19}$$

where $(M_t)_{t \ge 0}$ is a (local) martingale. Using the optional sampling theorem ($E_x(M_\tau) = 0$) this establishes a basic connection between X and $\mathbb{L}_X$, and both (1.1.9+1.1.10) and (1.1.17+1.1.18) are easily understood. This reasoning extends to more general Markov processes too.

The *Dirichlet condition* (1.1.18) alone is never sufficient to determine the solution uniquely. A crucial fact is that in many cases the following *Neumann condition* holds as well:

$$\frac{\partial V}{\partial x} = \frac{\partial M}{\partial x} \quad \text{on} \ \ \partial C \ . \tag{1.1.20}$$

This condition is termed *the principle of smooth-fit*. In one dimension it determines the solution uniquely, in higher dimension this may not be the case. We shall see later in this thesis that this condition fails for some Markov processes with jumps no matter how smooth M is.

The connection between optimal stopping and free-boundary problems (including the principle of smooth fit) was independently discovered by Mikhalevich [53], Chernoff [18], Lindley [48] and

McKean [52]. It is interesting that the thesis of Mikhalevich dates back to 1955. He studied the sequential testing problem for the mean of a Wiener process and apparently was the first to use the principle of smooth fit. This study has been taken further by Shiryaev [64]. McKean studied a free-boundary problem of parabolic type arising in optimal stopping of the American option pricing. No explicit formula has ever been found for the optimal stopping boundary in the finite horizon case. This work has been taken further by van Moerbeke [69] who studied its properties. A summary of the essential results on the American option pricing problem is given by Myneni [54]. A systematic study of analytic methods of optimal stopping problems by means of their reduction to free-boundary problems is initiated by Grigelionis and Shiryaev [37]. Beneš, Shepp and Witsenhausen [8] solve specific optimal control problems by reducing the Hamilton-Jacobi-Bellman equation to a free-boundary problem and using the principle of smooth fit. This connection between optimal control and optimal stopping goes back to Bather and Chernoff [4, 5]. A large number of papers have been written on this topic that has come to be called *singular stochastic control* (as optimal controls are local times).

Another class of problems by its very nature requires a procedure of sequential observation and consequently leads to optimal stopping and free-boundary problems. These are *quickest detection* problems. Their study was initiated by Shiryaev [62, 63]. A historical exposition of sequential testing problems and quickest detection problems is given in [65; pp.207-208]. Chapter 4 of that book contains the explicit solution of these problems in the case of a Wiener process.

Yet another class of optimal stopping problems deserves special mentioning as their methods are neither fully developed nor unified as yet. These problems appear under different names (the optimal selection problem, the best choice problem, the secretary problem, the house selling problem) but can commonly be called *optimal prediction* problems (see Section 2.3 in [65] and p.111 in the same book for references; see also the example in [26; pp.628-629]). In regard to problems of optimal prediction the papers [45, 34, 14, 36] are interesting to consult.

Free-boundary problems. Problems of free boundary originated in mathematical physics. A formal definition states that a boundary-value problem in a domain whose boundary is unknown is called a *free-boundary problem.* In addition to the standard boundary conditions needed to determine a solution, a new condition must be imposed at the free-boundary. One then seeks to determine both the function solving the equation and the free-boundary itself.

Two types of problems stimulated the development of the free-boundary theory. The first one is *the obstacle problem* (describing the shape of a membrane which is acted upon by an external force and constrained to remain above a solid obstacle) and *the Stefan problem* (describing the process of melting and solidification). More details on these problems can be found in a recent expository article by Friedman [30] where also several more recent free-boundary problems in science and technology are described (image development in electrography, chemical vapor deposition, coating flows, tumor growth). The article also contain other useful references on the subject that can help to trace the history. A list of older papers on free-boundary problems in the theory of differential equations can be found in [37].

Solutions to free-boundary problems for partial differential equations are rarely known explicitly. One is usually confined to study the existence, uniqueness, and asymptotic behavior of the solution. A melting ice problem presented in [29; Chapter 8] is a nice example.

1.2 A survey of the papers

Paper [P1] (Chapter 2): The solution is found to the optimal stopping problem with payoff:

$$\sup_{\tau} E\Big(S_\tau - \int_0^\tau c(X_t)\, dt \Big) \tag{1.2.1}$$

where $S = (S_t)_{t \ge 0}$ is the maximum process associated with the one-dimensional time-homogeneous diffusion $X = (X_t)_{t \ge 0}$, the function $x \mapsto c(x)$ is positive and continuous, and the supremum is taken over all stopping times τ of X for which the integral has finite expectation. It is proved, under no extra conditions, that this problem has a solution, i.e. the payoff is finite and there is an optimal stopping time, if and only if the following *maximality principle* holds: The first-order nonlinear differential equation

$$g'(s) = \frac{\sigma^2\big(g(s)\big)\, L'\big(g(s)\big)}{2\, c\big(g(s)\big) \Big(L(s) - L\big(g(s)\big) \Big)} \tag{1.2.2}$$

admits a maximal solution $s \mapsto g_*(s)$ which stays strictly below the diagonal in $I\!R^2$. [In this equation $x \mapsto \sigma(x)$ is the diffusion coefficient and $x \mapsto L(x)$ the scale function of X.] In this case the following stopping time:

$$\tau_* = \inf\, \{\, t > 0 \mid X_t \le g_*(S_t)\, \} \tag{1.2.3}$$

is proved optimal, and explicit formulas for the payoff are given. The result has a large number of applications, and may be viewed as the cornerstone in a general treatment of the maximum process.

Paper [P2] (Chapter 3): Let $B = (B_t)_{0 \le t \le 1}$ be standard Brownian motion started at zero, and let $S_t = \max_{0 \le r \le t} B_r$ for $0 \le t \le 1$. Consider the optimal stopping problem:

$$V_* = \inf_{\tau} E\big(B_\tau - S_1\big)^2 \tag{1.2.4}$$

where the infimum is taken over all stopping times of B satisfying $0 \le \tau \le 1$. We show that the infimum is attained at the stopping time:

$$\tau_* = \inf\, \Big\{\, 0 \le t \le 1 \mid S_t - B_t \ge z_* \sqrt{1-t}\, \Big\} \tag{1.2.5}$$

where $z_* = 1.12\ldots$ is the unique root of the equation:

$$4\Phi(z_*) - 2z_*\varphi(z_*) - 3 = 0 \tag{1.2.6}$$

with $\varphi(x) = (1/\sqrt{2\pi})e^{-x^2/2}$ and $\Phi(x) = \int_{-\infty}^x \varphi(y)\, dy$. The value V_* equals $2\Phi(z_*) - 1$. The method of proof relies upon a stochastic integral representation of S_1 , time-change arguments, and the solution of a free-boundary problem.

Paper [P3] (Chapter 4): We present the explicit solution of the Bayesian and variational problem of sequential testing of two simple hypotheses about the intensity of an observed Poisson process. The method of proof consists of reducing the initial problem to a free-boundary differential-difference problem, and solving the latter by use of the principles of smooth and continuous fit.

A rigorous proof of the optimality of the Wald's sequential probability ratio test in the variational formulation of the problem is obtained as a consequence of the solution of the Bayesian problem.

Paper [**P4**] (Chapter 5): The Poisson disorder problem seeks to determine a stopping time which is as close as possible to the (unknown) time of 'disorder' when the intensity of an observed Poisson process changes from one value to another. Partial answers to this question are known to date only in some special cases, and the main purpose of the present paper is to describe the structure of the solution in the general case. The method of proof consists of reducing the initial (optimal stopping) problem to a free-boundary differential-difference problem. The key point in the solution is reached by specifying when the principle of smooth fit breaks down and gets superseded by the principle of continuous fit. This can be done in probabilistic terms (by describing the sample path behaviour of the a posteriori probability process) and in analytic terms (via the existence of a singularity point of the free-boundary equation).

Paper [**P5**]: If $B = (B_t)_{t \geq 0}$ is a standard Brownian motion started at x under P_x for $x \geq 0$, and τ is any stopping time for B with $E_x(\tau) < \infty$, then for each $p > 1$ the following inequality is shown to be sharp:

(1.2.7) $$E_x\Big(\max_{0 \leq t \leq \tau} |B_t|^p\Big) \leq \Big(\frac{p}{p-1}\Big)^p E_x|B_\tau|^p - \Big(\frac{p}{p-1}\Big) x^p \; .$$

The sharpness is realized through the stopping times of the form:

(1.2.8) $$\tau_{\lambda,\varepsilon} = \inf\Big\{\, t > 0 \;\Big|\; \max_{0 \leq s \leq t} B_s - \lambda B_t \geq \varepsilon \,\Big\}$$

for which it is computed:

(1.2.9) $$E_0\big(\tau_{\lambda,\varepsilon}\big) = \frac{\varepsilon^2}{\lambda(2-\lambda)}$$

whenever $\varepsilon > 0$ and $0 < \lambda < 2$. Hence, for the stopping time:

(1.2.10) $$\sigma_{\lambda,\varepsilon} = \inf\Big\{\, t > 0 \;\Big|\; \max_{0 \leq s \leq t} |B_s| - \lambda |B_t| \geq \varepsilon \,\Big\}$$

which is a convolution of $\tau_{\lambda,\lambda\varepsilon}$ with the first hitting time of ε by $|B| = (|B_t|)_{t \geq 0}$, we have:

(1.2.11) $$E_0\big(\sigma_{\lambda,\varepsilon}\big) = \frac{2\varepsilon^2}{(2-\lambda)}$$

for all $\varepsilon > 0$ and all $0 < \lambda < 2$. The method of proof relies upon the principle of smooth fit and the maximality principle for a free-boundary problem, and Itô-Tanaka's formula (being applied two-dimensionally). The main emphasis is on the explicit formulas throughout obtained.

Paper [**P6**]: Let $B = (B_t)_{t \geq 0}$ be standard Brownian motion started at zero. Then the following inequality is shown to be satisfied:

(1.2.12) $$E\Big(\max_{0 \leq t \leq \tau} |B_t|^p\Big) \leq \gamma^*_{p,q}\Big(E \int_0^\tau |B_t|^{q-1}\, dt\Big)^{p/(q+1)}$$

for all stopping times τ for B, all $0 < p < 1+q$, and all $q > 0$, with the best possible value for the constant being equal:

$$\gamma^*_{p,q} = (1+\kappa)\left(\frac{s_*}{\kappa^\kappa}\right)^{1/(1+\kappa)} \tag{1.2.13}$$

where $\kappa = p/(q-p+1)$, and s_* is the zero point of the (unique) maximal solution $s \mapsto g_*(s)$ of the following differential equation:

$$g^\alpha(s)\Big(s^\beta - g^\beta(s)\Big)\frac{dg}{ds}(s) = K \tag{1.2.14}$$

satisfying $0 < g_*(s) < s$ for all $s > s_*$, where $\alpha = q/p - 1$, $\beta = 1/p$ and $K = p/2$. This solution is also characterized by $g_*(s)/s \to 1$ for $s \to \infty$. The equality above is attained at the following stopping time:

$$\tau_* = \inf\{\, t > 0 \mid X_t = g_*(S_t)\,\} \tag{1.2.15}$$

where $X_t = |B_t|^p$ and $S_t = \max_{0 \le r \le t} |B_r|^p$. In the case $p=1$ the closed form for $s \mapsto g_*(s)$ is found. This yields $\gamma^*_{1,q} = (q(q+1)/2)^{1/(q+1)}(\Gamma(1+(q+1)/q))^{q/(q+1)}$ for all $q > 0$. In the case $p \neq 1$ no closed form for $s \mapsto g_*(s)$ seems to exist. The inequality above holds also in the case $p = q+1$ (Doob's maximal inequality). In this case the equation above (with $K = p/2c$) admits $g_*(s) = \lambda s$ as the maximal solution, and the equality is attained only in the limit through the stopping times $\tau_* = \tau^*(c)$ when c tends to the best value $\gamma^*_{q+1,q} = (q+1)^{q+2}/2q^q$ from above. The method of proof relies upon the principle of smooth fit and the maximality principle. The results obtained extend to the case when B starts at any given point, as well as to all non-negative submartingales.

Paper [P7]: Let $B = (B_t)_{t \ge 0}$ be standard Brownian motion started at zero. We prove:

$$E\Big(\max_{0 \le t \le \tau} |B_t|\Big) \le 1 + \frac{1}{e^c(c-1)} + cE\Big(|B_\tau| \log^+ |B_\tau|\Big) \tag{1.2.16}$$

for all $c > 1$ and all stopping times τ for B satisfying $E(\tau^r) < \infty$ for some $r > 1/2$. This inequality is sharp, and the equality is attained at the stopping time:

$$\tau_* = \inf\ \big\{\, t > 0 \mid S_t \ge u_*\,,\ X_t = 1 \vee \alpha S_t\,\big\} \tag{1.2.17}$$

where $u_* = 1+1/e^c(c-1)$ and $\alpha = (c-1)/c$ for $c > 1$ with $X_t = |B_t|$ and $S_t = \max_{0 \le r \le t} |B_r|$. Likewise, we prove the following inequality:

$$E\Big(\max_{0 \le t \le \tau} |B_t|\Big) \le \frac{c^2}{e(c-1)} + cE\Big(|B_\tau| \log |B_\tau|\Big) \tag{1.2.18}$$

for all $c > 1$ and all stopping times τ for B satisfying $E(\tau^r) < \infty$ for some $r > 1/2$. This inequality is sharp, and the equality is attained at the stopping time:

$$\sigma_* = \inf\ \big\{\, t > 0 \mid S_t \ge v_*\,,\ X_t = \alpha S_t\,\big\} \tag{1.2.19}$$

where $v_* = c/e(c-1)$ and $\alpha = (c-1)/c$ for $c > 1$. These results contain and refine the results on the $L \log L$-inequality of Gilat [32] which are obtained by analytic methods. The method of proof given here is probabilistic and is based upon solving the optimal stopping problem with the payoff:

$$V = \sup_{\tau} E\big(S_\tau - cF(X_\tau)\big) \tag{1.2.20}$$

where $F(x)$ equals either $x \log^+ x$ or $x \log x$. This optimal stopping problem has some new interesting features, but in essence is solved by applying the principle of smooth fit and the maximality principle. The results extend to the case when B starts at any given point (as well as to all non-negative submartingales).

Paper [P8]: Let $B = (B_t)_{t \ge 0}$ be standard Brownian motion started at zero. Then the following inequality is shown to be satisfied:

$$E\Big(\max_{0 \le t \le \tau} |B_t|\Big) \le cE\Big(\int_0^\tau \frac{dt}{1+|B_t|}\Big) + \frac{1}{2c-1} \tag{1.2.21}$$

for all stopping times τ for B and all $c > 1/2$. The stopping times at which the equality is attained are of the form:

$$\tau_c = \inf\ \big\{\, t > 0 \ \mid\ S_t - \alpha X_t \ge \beta \,\big\} \tag{1.2.22}$$

where $\alpha = 1 + 1/(2c-1)$, $\beta = 1/(2c-1)$, $X_t = |B_t|$ and $S_t = \max_{0 \le r \le t} |B_r|$. Taking infimum over all $c > 1/2$ we obtain:

$$E\Big(\max_{0 \le t \le \tau} |B_t|\Big) \le \frac{1}{2} E\Big(\int_0^\tau \frac{dt}{1+|B_t|}\Big) + \sqrt{2}\Big(E\int_0^\tau \frac{dt}{1+|B_t|}\Big)^{1/2} \tag{1.2.23}$$

for all stopping times τ for B . This inequality is sharp (the equality is attained at each τ_c for all $c > 1/2$). In view of Itô-Tanaka's formula these inequalities may be thought of as the integral analogues (for reflected Brownian motion) of the classical $L \log L$-inequality of Hardy and Littlewood. The proof is based upon solving the optimal stopping problem:

$$V = \sup_{\tau} E\big(S_\tau - cI_\tau\big) \tag{1.2.24}$$

where $I_\tau = \int_0^\tau (1+|B_t|)^{-1} dt$. The payoff V is shown to be finite if and only if $c > 1/2$, and in this case $V = 1/(2c-1)$. The optimal stopping problem is solved by applying the principle of smooth fit and the maximality principle. All results extend to the case when Brownian motion B starts at any given point.

Paper [P9]: Motivated by applications in option pricing theory we formulate and solve the following problem. Given a standard Brownian motion $B = (B_t)_{t \ge 0}$ and a centered probability measure μ on $I\!R$ having the distribution function F with a strictly positive density F' satisfying:

$$\int_0^\infty x \log x \ \mu(dx) < \infty \tag{1.2.25}$$

there exists a cost function $x \mapsto c(x)$ in the optimal stopping problem:

$$\sup_{\tau} E\bigg(\max_{0\le t\le\tau} B_t - \int_0^{\tau} c(B_t)\, dt \bigg) \tag{1.2.26}$$

such that for the optimal stopping time τ_* we have:

$$B_{\tau_*} \sim \mu \; . \tag{1.2.27}$$

The cost function is explicitly given by the formula:

$$c(x) = \frac{1}{2} \frac{F'(x)}{\big(1-F(x)\big)} \tag{1.2.28}$$

where one incidentally recognizes $x \mapsto F'(x)/(1-F(x))$ as *the Hazard function* of μ . There is also a simple explicit formula for the optimal stopping time τ_* , but the main emphasis of the result is on the existence of the underlying functional in the optimal stopping problem. The integrability condition on μ is natural and cannot be improved. The condition on the existence of a strictly positive density is imposed for simplicity, and more general cases could be treated similarly. The method of proof combines ideas and facts on optimal stopping of the maximum process [P1] and the Azéma-Yor solution of the Skorokhod-embedding problem [2, 3]. A natural connection between these two theories is established, and new facts of interest for both are proved. The result extends in a similar form to stochastic integrals with respect to B , as well as to more general diffusions driven by B .

Paper [P10]: In the context of the Russian option we present a new derivation of the solution for the underlying one-dimensional optimal stopping problem. Our method is not based on the smooth pasting guess and Itô formula, but only uses the strong Markov property. In addition, the exact formula is given for the expected waiting time of the optimal stopping strategy. Two different methods for this computation are presented. Both methods can be easily generalized to treat similar problems for general one-dimensional time-homogeneous diffusions.

Paper [P11]: Explicit formulas are found for the payoff and the optimal stopping strategy of the following optimal stopping problem:

$$\sup_{\tau} E\bigg(\max_{0\le t\le\tau} X_t - c\tau \bigg) \tag{1.2.29}$$

where $X = (X_t)_{t\ge 0}$ is geometric Brownian motion with drift μ and volatility $\sigma > 0$, and the supremum is taken over all stopping times for X . The payoff is shown to be finite, if and only if $\mu < 0$. The optimal stopping time is given by:

$$\tau_* = \inf \bigg\{ t > 0 \mid X_t = g_* \bigg(\max_{0\le s\le t} X_s \bigg) \bigg\} \tag{1.2.30}$$

where $s \mapsto g_*(s)$ is the *maximal* solution of the (nonlinear) differential equation:

$$\frac{\partial g}{\partial s} = K \frac{g^{\Delta+1}}{s^{\Delta} - g^{\Delta}} \quad (s > 0) \tag{1.2.31}$$

under the condition $0 < g(s) < s$, where $\Delta = 1 - 2\mu/\sigma^2$ and $K = \Delta\sigma^2/2c$. The following estimate is established:

$$g_*(s) \sim \left(\frac{\Delta-1}{K\Delta}\right)^{1/\Delta} s^{1-1/\Delta} \tag{1.2.32}$$

as $s \to \infty$. Applying these results we prove the following maximal inequality:

$$E\left(\max_{0\le t\le\tau} X_t\right) \le 1 - \frac{\sigma^2}{2\mu} + \frac{\sigma^2}{2\mu}\exp\left(-\frac{(\sigma^2-2\mu)^2}{2\sigma^2}E(\tau) - 1\right) \tag{1.2.33}$$

where τ may be any stopping time for X. This extends the well-known identity:

$$E\left(\sup_{t>0} X_t\right) = 1 - \frac{\sigma^2}{2\mu} \tag{1.2.34}$$

and is shown to be sharp. The method of proof relies upon a smooth pasting guess (for a free-boundary problem) and Itô-Tanaka's formula (being applied two-dimensionally). The key point and main novelty in our approach is the maximality principle for the moving boundary (the optimal stopping boundary is the maximal solution of the differential equation obtained by a smooth pasting guess). We think that this principle is by itself of theoretical and practical interest.

Paper [**P12**]: Given a one-dimensional diffusion the solution is found to the optimal stopping problem where the gain is given by the maximum of the process and the cost is proportional to the duration of time. The optimal stopping boundary is shown to be the maximal solution of a nonlinear differential equations expressed in terms of the scale function and the speed measure. Applications to maximal inequalities are indicated.

Paper [**P13**]: Some non-linear optimal stopping problems can be solved explicitly by using a common method which is based on time-change. We describe this method and illustrate its use by considering several examples dealing with Brownian motion. In each of these examples we derive explicit formulas for the value function and display the optimal stopping time. The main emphasis of the paper is on the method of proof and its unifying scope.

Paper [**P14**]: Let $B = (B_t)_{t\ge0}$ be a Brownian motion started at $x \in I\!R$. Given a stopping time τ for B and a real valued map F, we show how one can optimally bound:

$$E\left(\int_0^\tau F\big(|B_t|\big)\,dt\right) \tag{1.2.35}$$

in terms of $E(\tau)$. The method of proof relies upon solving the optimal stopping problem where one minimizes or maximizes:

$$E\left(\int_0^\tau F\big(|B_t|\big)\,dt - c\tau\right) \tag{1.2.36}$$

over all stopping times τ for B with $c > 0$. By Itô's formula and the optional sampling theorem this problem simplifies to the form where explicit computations are possible. The method is quantitatively demonstrated through the example of $F(x) = x^r$ with $r > -1$. This yields some new sharp inequalities.

Paper [**P15**]: The solution is presented to all optimal stopping problems of the form:

$$\sup_{\tau} E\Big(G\Big(|B_\tau|\Big) - c\tau\Big) \tag{1.2.37}$$

where $B = (B_t)_{t\ge 0}$ is standard Brownian motion and the supremum is taken over all stopping times τ for B with finite expectation, while the map $G : I\!R_+ \to I\!R$ satisfies $G\big(|x|\big) \le c|x|^2 + d$ for some $d \in I\!R$ with $c > 0$ being given and fixed. The optimal stopping time is shown to be the hitting time by the reflecting Brownian motion $|B| = (|B_t|)_{t\ge 0}$ of the set of all (approximate) maximum points of the map $x \mapsto G(|x|) - cx^2$. The method of proof relies upon Wald's identity for Brownian motion and simple real analysis arguments. A simple proof of the Dubins-Jacka-Schwarz-Shepp-Shiryaev (square root of two) maximal inequality for randomly stopped Brownian motion is given as an application.

Paper [**P16**]: Let $B = (B_t)_{t\ge 0}$ be standard Brownian motion started at 0 under P , let $S_t = \max_{0\le r\le t} B_r$ be the maximum process associated with B , and let $g : I\!R_+ \to I\!R$ be a (strictly) monotone continuous function satisfying $g(s) < s$ for all $s \ge 0$. Let τ be the first-passage time of B over $t \mapsto g(S_t)$:

$$\tau = \inf\,\big\{\, t > 0 \mid B_t \le g(S_t) \,\big\} \; . \tag{1.2.38}$$

Let G be the function defined by:

$$G(y) = \exp\bigg(- \int_0^{g^{-1}(y)} \frac{ds}{s - g(s)} \bigg) \tag{1.2.39}$$

for $y \in I\!R$ in the range of g . Then, if g is increasing, we have:

$$\lim_{t\to\infty} \sqrt{t}\; P\{\, \tau \ge t \,\} = \sqrt{\frac{2}{\pi}} \bigg(-g(0) - \int_{g(0)}^{g(\infty)} G(y)\; dy \bigg) \tag{1.2.40}$$

and this number is finite. Similarly, if g is decreasing, we have:

$$\lim_{t\to\infty} \sqrt{t}\; P\{\, \tau \ge t \,\} = \sqrt{\frac{2}{\pi}} \bigg(-g(0) + \int_{g(\infty)}^{g(0)} G(y)\; dy \bigg) \tag{1.2.41}$$

and this number may be infinite. These results may be viewed as a *stochastic boundary* extension of some known results on the first-passage time over deterministic boundaries. The method of proof relies on the classical Tauberian theorem and certain extensions of the Novikov-Kazamaki criteria for exponential martingales.

Paper [**P17**]: Given the maximum process $(S_t)_{t\ge 0} = (\max_{0\le r\le t} X_r)_{t\ge 0}$ associated with a diffusion $((X_t)_{t\ge 0}, P_x)$, and a continuous function g satisfying $g(s) < s$, we show how to compute the expectation of the Azéma-Yor stopping time:

$$\tau_g = \inf\,\{\, t > 0 \mid X_t \le g(S_t) \,\} \tag{1.2.42}$$

as a function of x . The method of proof is based upon verifying that the expectation solves a differential equation with two boundary conditions. The third 'missing' condition is formulated in the form of a minimality principle which states that the expectation is the minimal non-negative solution to this system. It enables us to express this solution in a closed form. The result is applied in the case when $(X_t)_{t\geq 0}$ is a Bessel process and g is a linear function.

Paper [P18]: Let $B = (B_t)_{t\geq 0}$ be standard Brownian motion started at zero, let $\mu > 0$ be given and fixed, and let ν be a probability measure on $I\!R$ having a strictly positive density F'. Then there exists a stopping time τ_* of B such that:

$$(B_{\tau_*}+\mu\tau_*) \sim \nu \tag{1.2.43}$$

if and only if the following condition is satisfied:

$$D_\mu := \int_{I\!R} e^{-2\mu x}\, \nu(dx) \leq 1\ . \tag{1.2.44}$$

Setting in this case $C_\mu = -(2\mu)^{-1}\log(D_\mu)$, the following *explicit* formula is valid:

$$\tau_* = \inf\Big\{\, t>0 \;\Big|\; (B_t+\mu t) \leq h_\mu\Big(\max_{0\leq r\leq t}(B_r+\mu r)\Big)\Big\} \tag{1.2.45}$$

where the map $s \mapsto h_\mu(s)$ for $s > C_\mu$ is expressed through its inverse by:

$$h_\mu^{-1}(x) = -\frac{1}{2\mu}\log\Big(\frac{1}{1-F(x)}\int_x^\infty e^{-2\mu t}\, dF(t)\Big) \quad (x \in I\!R) \tag{1.2.46}$$

and we set $h_\mu(s) = -\infty$ for $s \leq C_\mu$. This settles the question raised in a paper. In addition, it is proved that τ_* is pointwise the smallest possible stopping time satisfying $(B_{\tau_*}+\mu\tau_*) \sim \nu$ which generates stochastically the largest possible maximum of the process $(B_t+\mu t)_{t\geq 0}$ up to the time of stopping. This *minimax* property characterizes τ_* uniquely. The result recovers the Azéma-Yor solution of the Skorokhod embedding problem [2] by passing to the limit when $\mu \downarrow 0$. The condition on the existence of a strictly positive density is imposed for simplicity, and more general cases can be treated similarly. The line of arguments used in the proof can be extended to treat the case of more general *nonrecurrent* diffusions.

Paper [P19]: Let $X = (X_t)_{t\geq 0}$ be a one-dimensional time-homogeneous diffusion process associated with the infinitesimal operator:

$$I\!L_X = \mu(x)\,\frac{\partial}{\partial x} + \frac{\sigma^2(x)}{2}\,\frac{\partial^2}{\partial x^2} \tag{1.2.47}$$

where $x \mapsto \mu(x)$ and $x \mapsto \sigma(x) > 0$ are continuous. We show how the question of finding a function $x \mapsto H(x)$ such that:

$$c_1 E\big(H(\tau)\big) \leq E\Big(\max_{0\leq t\leq \tau}|X_t|\Big) \leq c_2 E\big(H(\tau)\big) \tag{1.2.48}$$

holds for all stopping times τ of X relates to solutions of the equation:

$$
\mathbb{L}_X(F) = 1 \ . \tag{1.2.49}
$$

Explicit expressions for H are derived in terms of μ and σ . The method of proof relies upon a domination principle established by Lenglart and Itô calculus.

Paper [**P20**]: Let $V = (V_t)_{t \geq 0}$ be the Ornstein-Uhlenbeck velocity process solving:

$$
dV_t = -\beta V_t \ dt + dB_t \tag{1.2.50}
$$

with $V_0 = 0$, where $\beta > 0$ and $B = (B_t)_{t \geq 0}$ is a standard Brownian motion. Then there exist universal constants $C_1 > 0$ and $C_2 > 0$ such that:

$$
\frac{C_1}{\sqrt{\beta}} E\sqrt{\log(1+\beta\tau)} \leq E\Big(\max_{0 \leq t \leq \tau} |V_t|\Big) \leq \frac{C_2}{\sqrt{\beta}} E\sqrt{\log(1+\beta\tau)} \tag{1.2.51}
$$

for all stopping times τ of V . In particular, this yields the existence of universal constants $D_1 > 0$ and $D_2 > 0$ such that:

$$
D_1 \, E\sqrt{\log\big(1+\log(1+\tau)\big)} \leq E\Big(\max_{0 \leq t \leq \tau} \frac{|B_t|}{\sqrt{1+t}}\Big) \leq D_2 \, E\sqrt{\log\big(1+\log(1+\tau)\big)} \tag{1.2.52}
$$

for all stopping times τ of B . This inequality may be viewed as a stopped law of iterated logarithm. The method of proof relies upon a variant of Lenglart's domination principle and makes use of Itô calculus.

Paper [**P21**]: Let $V = (V_t)_{t \geq 0}$ be the Ornstein-Uhlenbeck velocity process solving:

$$
dV_t = -\beta V_t \ dt + \sigma \, dB_t \tag{1.2.53}
$$

with $V_0 = 0$, where $\beta > 0$, $\sigma > 0$ and $B = (B_t)_{t \geq 0}$ is a standard Brownian motion. Then the following inequality is satisfied:

$$
E\Big(\max_{0 \leq t \leq \tau} |V_t|\Big) \leq s_*(\kappa; \beta, \sigma) + \frac{\kappa}{\beta} E\Big(e^{(\beta/\sigma^2) V_\tau^2} - 1\Big) \tag{1.2.54}
$$

for all bounded stopping times τ of V and all $\kappa > 0$, where $s_*(\kappa; \beta, \sigma) > 0$ is the unique zero point of the maximal solution of the equation:

$$
y' = \frac{\sigma^2}{2\kappa \displaystyle\int_y^x e^{(\beta/\sigma^2) z^2} \, dz} \tag{1.2.55}
$$

staying strictly below the diagonal in $\mathbb{R}^2$. This inequality is sharp, and equality can be attained for each $\kappa > 0$. The following estimate is established:

$$
s_*(\kappa; \beta, \sigma) \leq \frac{\sigma}{\sqrt{\beta}} \, \Psi^{-1}\Big(\frac{\sqrt{\beta}\,\sigma}{2\kappa}\Big) \tag{1.2.56}
$$

where $\Psi(x) = \int_0^x e^{z^2} dz$. In particular, this yields the existence of a universal constant $C \geq \sqrt{2}$

such that the following inequality holds:

$$E\Big(\max_{0\le t\le\tau}|V_t|\Big)\;\le\;C\,\frac{\sigma}{\sqrt{\beta}}\sqrt{\log E\Big(e^{(\beta/\sigma^2)V_\tau^2}\Big)} \tag{1.2.57}$$

for all stopping times τ of V for which $(e^{(\beta/\sigma^2)V_{\tau\wedge t}^2})_{t\ge0}$ is uniformly integrable. This inequality shows that the question of controlling the velocity of Brownian motion by its terminal value can be answered positively. Better versions of this inequality are also derived which go beyond the best value for C .

Paper [**P22**]: Let $\beta=(\beta_t)_{t\ge0}$ denote the unique strong solution of the equation:

$$d\beta_t\;=\;-\mu\;sign(\beta_t)\;dt+dB_t \tag{1.2.58}$$

satisfying $\beta_0=0$, where $\mu>0$ and $B=(B_t)_{t\ge0}$ is a standard Brownian motion. Then $|\beta|=(|\beta_t|)_{t\ge0}$ is known to be a realisation of the reflected Brownian motion with drift $-\mu$. Using this representation we show that there exist universal constants $c_1>0$ and $c_2>0$ such that:

$$c_1E\Big(H_\mu(\tau)\Big)\;\le\;E\Big(\max_{0\le t\le\tau}|\beta_t|\Big)\;\le\;c_2E\Big(H_\mu(\tau)\Big) \tag{1.2.59}$$

for all stopping times τ of β , where $H_\mu(x)=F_\mu^{-1}(x)$ denotes the inverse of the map $F_\mu(x)=(e^{2\mu x}-2\mu x-1)/2\mu^2$. In addition, we show that:

$$E\Big(\max_{0\le t\le\tau}|\beta_t|\Big)\;\le\;G_\mu\Big(E(\tau)\Big) \tag{1.2.60}$$

for all stopping times τ of β , where $G_\mu(x)=\inf_{c>0}(cx+(1/2\mu)\log(1+\mu/c))$. Both inequalities have their well-known analogues for Brownian motion (obtained by letting $\mu\downarrow0$). The method of proof relies upon Lenglart's domination principle, Itô calculus, and optimal stopping techniques.

2. Maximum process problems

Our main aim in this chapter is to present the solution to a problem of optimal stopping for the maximum process associated with a one-dimensional time-homogeneous diffusion. The solution found has a large number of applications, and may be viewed as the cornerstone in a general treatment of the maximum process.

2.1 Introduction

In the setting of (2.2.1)-(2.2.3) we consider the optimal stopping problem (2.2.4), where the supremum is taken over all stopping times τ satisfying (2.2.5), and the cost function c is positive and continuous. The main result of the chapter is presented in Theorem 2.3.1, where it is proved that this problem has a solution (the payoff is finite and there is an optimal stopping strategy) if and only if *the maximality principle* holds, i.e. the first-order nonlinear differential equation (2.3.21) admits a maximal solution which stays strictly below the diagonal in $\mathbb{R}^2$ (see *Figure* 2.1 below). The maximal solution is proved to be an optimal stopping boundary, i.e. the stopping time (2.3.31) is optimal, and the payoff is given explicitly by (2.3.30). Moreover, this stopping time is shown to be pointwise the smallest possible optimal stopping time. If there is no such maximal solution of (2.3.21), the payoff is proved to be infinite and there is no optimal stopping time. The chapter finishes with four examples in Section 2.4 which are aimed to illustrate some applications of the result proved.

The optimal stopping problem (2.2.4) has been considered in some special cases earlier. Jacka [40] treats the case of reflected Brownian motion, while Dubins, Shepp and Shiryaev [24] treat the case of Bessel processes. In these papers the problem was solved very effectively by guessing the nature of the optimal stopping boundary and making use of the principle of smooth-fit. The same is true for the "discounted" problem (2.3.60) with $c \equiv 0$ in the case of geometric Brownian motion which in the framework of option pricing theory (Russian option) was solved by Shepp and Shiryaev in [60] (see also [61] and [P10]). For the first time a strong need for additional arguments was felt in [P11], where the problem (2.2.4) for geometric Brownian motion was considered with the cost function $c(x) \equiv c > 0$. There, by use of Picard's method of successive approximations, it was proved that the maximal solution of (2.3.21) is an optimal stopping boundary, and since this solution could not be expressed in closed form, it really showed the full power of the method. Such non-trivial solutions were also obtained in [24] by a method which relies on estimates of the payoff obtained a priori. Motivated by similar ideas, sufficient conditions for the maximality principle to hold for general diffusions are given in [P12]. The method of proof used there relies on a transfinite induction argument. In order to solve the problem in general, the fundamental question was how to relate the maximality principle to the superharmonic characterization of the payoff, which is the key result in the general theory. This fact has been indicated by A. Shiryaev.

The most interesting point in our solution of the optimal stopping problem (2.2.4) relies on the fact that we have now described this connection, and actually proved that the maximality principle is equivalent to the superharmonic characterization of the payoff (for a three-dimensional process). The crucial observations in this direction are (2.3.28) and (2.3.29), which show that the only possible optimal stopping boundary is the maximal solution (see (2.3.38) in the proof of Theorem 2.3.1). In the next step of proving that the maximal solution is indeed an optimal stopping boundary, it was crucial to make use of so-called “bad-good” solutions of (2.3.21), “bad” in the sense that they hit the diagonal in $\mathbb{R}^2$, and “good” in the sense that they are not too large (see *Figure* 2.1 below). These “bad-good” solutions are used to approximate the maximal solution in a desired manner, see the proof of Theorem 2.3.1 (starting from (2.3.40) onwards), and this turns out to be the key argument in completing the proof.

Our methodology adopts and extends earlier results of Dubins, Shepp and Shiryaev [24], and is, in fact, quite standard in the business of solving particular optimal stopping problems: (i) one tries to guess the nature of the optimal stopping boundary as a member of a “reasonable” family; (ii) computes the expected reward; (iii) maximises this over the family; (iv) and then tries to argue that the resulting stopping time is optimal in general. This process is often facilitated by “ad hoc” principles, as the famous “principle of smooth-fit” for instance. This procedure is used very effectively in this chapter too, as opposed to results from the general theory of optimal stopping, and as suggested by the referee, we should like to stress this fact. We would also like to point out, however, that the maximality principle of the present chapter should rather be seen as a convenient reformulation of the basic principle on a superharmonic characterization from the general theory, than a new principle on its own. A. Shiryaev has also noticed a similar maximality property of his solution a long while ago [64; Figure 3], and that similar tricks were used by other people too; see also [P17] for a related result.

2.2 Formulation of the problem

Let $X = (X_t)_{t \geq 0}$ be a one-dimensional time-homogeneous *diffusion* process associated with the infinitesimal operator:

$$\mathbb{L}_X = \mu(x) \frac{\partial}{\partial x} + \frac{\sigma^2(x)}{2} \frac{\partial^2}{\partial x^2} \tag{2.2.1}$$

where the *drift* coefficient $x \mapsto \mu(x)$ and the *diffusion* coefficient $x \mapsto \sigma(x) > 0$ are continuous. Assume moreover that there exists a standard *Brownian motion* $B = (B_t)_{t \geq 0}$ defined on $(\Omega, \mathcal{F}, P)$ such that X solves the stochastic differential equation:

$$dX_t = \mu(X_t)\, dt + \sigma(X_t)\, dB_t \tag{2.2.2}$$

with $X_0 = x$ under $P_x := P$ for $x \in \mathbb{R}$. The state space of X is assumed to be $\mathbb{R}$.

With X we associate the maximum process:

$$S_t = \Big(\max_{0 \leq r \leq t} X_r \Big) \vee s \tag{2.2.3}$$

started at $s \geq x$ under $P_{x,s} := P$. The main objective of this chapter is to present the solution to the optimal stopping problem with payoff:

$$V_*(x,s) = \sup_\tau E_{x,s} \Big(S_\tau - \int_0^\tau c(X_t)\, dt \Big) \tag{2.2.4}$$

where the supremum is taken over stopping times τ of X satisfying:

$$E_{x,s}\Big(\int_0^\tau c(X_t)\,dt \Big) \;<\; \infty \tag{2.2.5}$$

and the *cost* function $x \mapsto c(x) > 0$ is continuous.

1. To state and prove the initial observation about (2.2.4), and for further reference, we need to recall a few general facts about one-dimensional diffusions (see [58; pp.270-303]).

The *scale function* of X is given by:

$$L(x) = \int^x \exp\Big(-\int^y \frac{2\mu(z)}{\sigma^2(z)}\;dz\Big)\;dy \tag{2.2.6}$$

for $x \in I\!R$. Throughout we denote:

$$\tau_x \,=\, \inf\,\{\, t > 0 \;|\; X_t \,=\, x \,\} \tag{2.2.7}$$

and set $\tau_{x,y} \,=\, \tau_x \wedge \tau_y$. Then we have:

$$P_x\big(X_{\tau_{a,b}} = a\big) \,=\, \frac{L(b)-L(x)}{L(b)-L(a)} \tag{2.2.8}$$

$$P_x\big(X_{\tau_{a,b}} = b\big) \,=\, \frac{L(x)-L(a)}{L(b)-L(a)} \tag{2.2.9}$$

whenever $a \,\le\, x \,\le\, b$.

The *speed measure* of X is given by:

$$m(dx) \,=\, \frac{2\,dx}{L'(x)\,\sigma^2(x)}\;. \tag{2.2.10}$$

The *Green function* of X on $[a,b]$ is defined by:

$$\begin{aligned} G_{a,b}(x,y) &= \frac{\big(L(b)-L(x)\big)\big(L(y)-L(a)\big)}{\big(L(b)-L(a)\big)} \quad \text{if} \;\; a \le y \le x \\ &= \frac{\big(L(b)-L(y)\big)\big(L(x)-L(a)\big)}{\big(L(b)-L(a)\big)} \quad \text{if} \;\; x \le y \le b\;. \end{aligned} \tag{2.2.11}$$

If $f \,:\, I\!R \mapsto I\!R$ is a measurable function, then:

$$E_x\Big(\int_0^{\tau_{a,b}} f(X_t)\;dt\Big) \,=\, \int_a^b f(y)\,G_{a,b}(x,y)\;m(dy)\;. \tag{2.2.12}$$

2. Due to the specific form of the optimal stopping problem (2.2.4), the following observation is nearly evident (see [24; pp.237-238]).

Proposition 2.2.1
The process $\bar{X}_t = (X_t, S_t)$ cannot be optimally stopped on the diagonal of $I\!R^2$.

Proof. Fix $x \in I\!R$, and set $l_n = x - 1/n$ and $r_n = x + 1/n$. Denoting $\tau_n = \tau_{l_n,r_n}$ it will be enough to show that:

$$E_{x,x}\Big(S_{\tau_n} - \int_0^{\tau_n} c(X_t)\,dt\Big) > x \tag{2.2.13}$$

for $n \ge 1$ large enough.

For this, note first by the strong Markov property and (2.2.8)-(2.2.9) that:

$$\begin{aligned} E_{x,x}\big(S_{\tau_n}\big) &\ge x\,P_x\big(X_{\tau_n} = l_n\big) + r_n\,P_x\big(X_{\tau_n} = r_n\big) \\ &= x\,\frac{L(r_n)-L(x)}{L(r_n)-L(l_n)} + r_n\,\frac{L(x)-L(l_n)}{L(r_n)-L(l_n)} \\ &= x + (r_n - x)\,\frac{L(x)-L(l_n)}{L(r_n)-L(l_n)} = x + (r_n - x)\,\frac{L'(\xi_n)(x-l_n)}{L'(\eta_n)(r_n-l_n)} \ge x + K/n \end{aligned} \tag{2.2.14}$$

since $L \in C^1$. On the other hand $K_1 := \sup_{l_n \le z \le r_n} c(z) < \infty$. Thus by (2.2.10)-(2.2.12) we get:

$$\begin{aligned} E_{x,x}\Big(\int_0^{\tau_n} c(X_t)\,dt\Big) &\le K_1\,E_x(\tau_n) = 2K_1 \int_{l_n}^{r_n} G_{a,b}(x,y)\,\frac{dy}{\sigma^2(y)L'(y)} \\ &\le K_2\Big(\int_{l_n}^{x} \big(L(y)-L(l_n)\big)\,dy + \int_x^{r_n} \big(L(r_n)-L(y)\big)\,dy\Big) \\ &\le K_3\Big((x-l_n)^2 + (r_n-x)^2\Big) = 2K_3/n^2 \end{aligned} \tag{2.2.15}$$

since σ is continuous and $L \in C^1$. Combining (2.2.14) and (2.2.15) we clearly obtain (2.2.13) for $n \ge 1$ large enough. The proof is complete. □

3. For a survey and the definitive results of Engelbert and Schmidt on existence, uniqueness, and various other aspects of solutions of one-dimensional stochastic differential equations we refer to [43] (Chapter 5).

2.3 Optimal stopping of the maximum process

In the setting of (2.2.1)-(2.2.3) consider the optimal stopping problem (2.2.4) where the supremum is taken over all stopping times τ of X satisfying (2.2.5). Our main aim in this section is to present the solution to this problem (Theorem 2.3.1). We begin our exposition with a few observations on the underlying structure of (2.2.4) with a view to the Markovian theory of optimal stopping.

1. Note that $\bar{X}_t = (X_t, S_t)$ is a two-dimensional Markov process with the state space $D = \{\,(x,s) \in I\!R^2 \mid x \le s\,\}$, which can change (increase) in the second coordinate only after hitting the diagonal $x = s$ in $I\!R^2$. Off the diagonal, the process $\bar{X} = (\bar{X}_t)_{t\ge 0}$ changes only in the first coordinate and may be identified with X. Due to its form and behaviour at the diagonal, we claim that the infinitesimal operator of $\bar{X}$ may thus be formally described as follows:

$$\begin{aligned} &I\!L_{\bar{X}} = I\!L_X \quad \text{in} \quad x < s \\ &\frac{\partial}{\partial s} = 0 \quad \text{at} \quad x = s \end{aligned} \tag{2.3.1}$$

with $I\!L_X$ as in (2.2.1). This means that the infinitesimal operator of $\bar{X}$ is acting on a space of

C^2-functions f on D satisfying $(\partial f/\partial s)(s,s)=0$. Observe that we do not tend to specify the domain of $I\!\!L_{\bar{X}}$ precisely, but will only verify that if $f:D\rightarrow I\!\!R$ is a C^2-function which belongs to the domain, then $(\partial f/\partial s)(s,s)$ must be zero.

To see this, we shall apply Itô's formula to the process $f(X_t,S_t)$ and take the expectation under $P_{s,s}$. By the optional sampling theorem being applied to the continuous local martingale which appears in this process (localized if needed), we obtain:

$$
\begin{aligned}
(2.3.2)\qquad & \frac{E_{s,s}\big(f(X_t,S_t)\big)-f(s,s)}{t} = E_{s,s}\bigg(\frac{1}{t}\int_0^t (I\!\!L_X f)(X_r,S_r)\;dr\bigg) \\
&+ E_{s,s}\bigg(\frac{1}{t}\int_0^t \frac{\partial f}{\partial s}(X_r,S_r)\;dS_r\bigg) \;\longrightarrow\; I\!\!L_X f(s,s) \;+\; \frac{\partial f}{\partial s}(s,s)\bigg(\lim_{t\downarrow 0}\frac{E_{s,s}(S_t-s)}{t}\bigg)
\end{aligned}
$$

as $t\downarrow 0$. Due to $\sigma>0$, we have $t^{-1}E_{s,s}(S_t-s)\rightarrow\infty$ as $t\downarrow 0$, and therefore the limit above is infinite, unless $(\partial f/\partial s)(s,s)=0$. This completes the claim (see also [24; pp.238-239]).

2. The problem (2.2.4) can be related to the Markovian theory of optimal stopping by introducing the following functional:

$$
(2.3.3)\qquad A_t \,=\, a + \int_0^t c(X_r)\;dr
$$

with $a\ge 0$ given and fixed, and noting that $Z_t=(A_t,X_t,S_t)$ is a Markov process which starts at (a,x,s) under P . Its infinitesimal operator is obtained by adding $c(x)(\partial/\partial a)$ to the infinitesimal operator of $\bar{X}$, which combined with (2.3.1) leads to the formal description:

$$
\begin{aligned}
(2.3.4)\qquad & I\!\!L_Z \,=\, c(x)(\partial/\partial a)+I\!\!L_X \quad \text{in} \quad x<s \\
& \tfrac{\partial}{\partial s} \,=\, 0 \quad \text{at} \quad x=s
\end{aligned}
$$

with $I\!\!L_X$ as in (2.2.1). Given $Z=(Z_t)_{t\ge 0}$, introduce the *gain function* $G(a,x,s)=s-a$, note that the payoff (2.2.4) viewed in terms of the general theory ought to be defined as:

$$
(2.3.5)\qquad \widetilde{V}_*(a,x,s) \,=\, \sup_\tau E\big(G(Z_\tau)\big)
$$

where τ is a stopping time of Z satisfying $E(A_\tau)<\infty$, and observe that:

$$
(2.3.6)\qquad \widetilde{V}_*(a,x,s) \,=\, V_*(x,s)-a\ .
$$

This identity is the main reason that we abandon the general formulation (2.3.5) and simplify it to the form (2.2.4), and that we speak of optimal stopping for the process $\bar{X}_t=(X_t,S_t)$ rather than the process $Z_t=(A_t,X_t,S_t)$.

Let us point out that the contents of this subsection are used in the sequel merely to clarify the result and method in terms of the general theory.

3. From now on our main aim will be to show that the problem (2.2.4) reduces to the problem of solving a first-order nonlinear differential equation (for the optimal stopping boundary). To derive this equation we shall first try to get a feeling for the points in the state space $\{\,(x,s)\in I\!\!R^2 \mid x\le s\,\}$ at which the process $\bar{X}_t=(X_t,S_t)$ can be optimally stopped.

When on the vertical level s, the process $\bar{X}_t = (X_t, S_t)$ stays at the same level until it hits the diagonal $x = s$ in $I\!R^2$. During that time $\bar{X}$ does not change (increase) in the second coordinate. Due to the strictly positive cost in (2.2.4), it is clear that we should not let the process $\bar{X}$ run to much to the left, since it could be "too expensive" to get back to the diagonal in order to offset the "cost" spent to travel all that way. More specifically, given s there should exist a point $g_*(s) \leq s$ such that if the process (X, S) reaches the point $(g_*(s), s)$ we should stop it instantly. In other words, the stopping time:

$$\tau_* = \inf \{ \, t > 0 \mid X_t \leq g_*(S_t) \, \} \tag{2.3.7}$$

should be optimal for the problem (2.2.4). For this reason we call $s \mapsto g_*(s)$ an *optimal stopping boundary*, and our aim will be to prove its existence and to characterise it. Observe by Proposition 2.2.1 that we must have $g_*(s) < s$ for all s, and that $V_*(x, s) = s$ for all $x \leq g_*(s)$.

4. To compute the payoff $V_*(x, s)$ for $g_*(s) < x \leq s$, and to find the optimal stopping boundary $s \mapsto g_*(s)$, we are led to formulate the following system:

$$(I\!L_X V)(x, s) = c(x) \quad \text{for } g(s) < x < s \text{ with } s \text{ fixed} \tag{2.3.8}$$

$$\frac{\partial V}{\partial s}(x, s)\Big|_{x=s-} = 0 \quad (\textit{normal reflection}) \tag{2.3.9}$$

$$V(x, s)\Big|_{x=g(s)+} = s \quad (\textit{instantaneous stopping}) \tag{2.3.10}$$

$$\frac{\partial V}{\partial x}(x, s)\Big|_{x=g(s)+} = 0 \quad (\textit{smooth-fit}) \tag{2.3.11}$$

with $I\!L_X$ as in (2.2.1). Note that (2.3.8)+(2.3.9) are in accordance with the general theory upon using (2.3.4) and (2.3.6) above: the infinitesimal operator of the process being applied to the payoff must be zero in the continuation region. The condition (2.3.10) is evident. The condition (2.3.11) is not part of the general theory; it is imposed since we believe that in the "smooth" setting of the problem (2.2.4) the principle of smooth-fit should hold. This belief will be vindicated after the fact, when we show in Theorem 2.3.1, that the solution of the system (2.3.8)-(2.3.11) leads to the payoff of (2.2.4). The system (2.3.8)-(2.3.11) constitutes a *free-boundary problem* (see [65; pp.157-162]). It has been derived for the first time by Dubins, Shepp and Shiryaev [24] in the case of Bessel processes.

5. To solve the system (2.3.8)-(2.3.11) we shall consider a stopping time of the form:

$$\tau_g = \inf \{ \, t > 0 \mid X_t \leq g(S_t) \, \} \tag{2.3.12}$$

and the map:

$$V_g(x, s) = E_{x,s}\left(S_{\tau_g} - \int_0^{\tau_g} c(X_t) \; dt \right) \tag{2.3.13}$$

associated with it, where $s \mapsto g(s)$ is a given function such that both $E_{x,s}(S_{\tau_g})$ and $E_{x,s}(\int_0^{\tau_g} c(X_t) \; dt)$ are finite. Set $V_g(s) := V_g(s, s)$ for all s. Considering $\tau_{g(s),s} = \inf \{ \, t > 0 \mid X_t \notin]g(s), s[\, \}$ and using the strong Markov property of X at $\tau_{g(s),s}$, by (2.2.8)-(2.2.12) we find:

$$(2.3.14)\quad V_g(x,s) = s\,\frac{L(s)-L(x)}{L(s)-L(g(s))} + V_g(s)\,\frac{L(x)-L(g(s))}{L(s)-L(g(s))} - \int_{g(s)}^{s} G_{g(s),s}(x,y)c(y)\ m(dy)$$

for all $g(s) < x < s$.

In order to determine $V_g(s)$, we shall rewrite (2.3.14) as follows:

$$(2.3.15)\quad V_g(s) - s = \frac{L(s)-L(g(s))}{L(x)-L(g(s))}\left(\big(V_g(x,s)-s\big) + \int_{g(s)}^{s} G_{g(s),s}(x,y)c(y)\ m(dy)\right)$$

and then divide and multiply through by $x-g(s)$ to obtain:

$$(2.3.16)\qquad \lim_{x\downarrow g(s)} \frac{V_g(x,s)-s}{L(x)-L(g(s))} = \frac{1}{L'(g(s))}\,\frac{\partial V_g}{\partial x}(x,s)\Big|_{x=g(s)+} .$$

It is easily seen by (2.2.11) that:

$$(2.3.17)\quad \lim_{x\downarrow g(s)} \frac{L(s)-L(g(s))}{L(x)-L(g(s))}\int_{g(s)}^{s} G_{g(s),s}(x,y)c(y)\ m(dy) = \int_{g(s)}^{s}\big(L(s)-L(y)\big)c(y)\ m(dy) .$$

Thus, if *the condition of smooth-fit*:

$$(2.3.18)\qquad \frac{\partial V_g}{\partial x}(x,s)\Big|_{x=g(s)+} = 0$$

is satisfied, we see from (2.3.15)-(2.3.17) that the following identity holds:

$$(2.3.19)\qquad V_g(s) = s + \int_{g(s)}^{s}\big(L(s)-L(y)\big)c(y)\ m(dy) .$$

Inserting this into (2.3.14), and using (2.2.11)-(2.2.12), we get:

$$(2.3.20)\qquad V_g(x,s) = s + \int_{g(s)}^{x}\big(L(x)-L(y)\big)c(y)\ m(dy)$$

for all $g(s) \le x \le s$.

If we now forget the origin of $V_g(x,s)$ in (2.3.13), and consider it purely as defined by (2.3.20), then it is straightforward to verify that $(x,s)\mapsto V_g(x,s)$ solves the system (2.3.8)-(2.3.11) in the region $g(s)<x<s$ if and only if the C^1-function $s\mapsto g(s)$ solves the following first-order nonlinear differential equation:

$$(2.3.21)\qquad g'(s) = \frac{\sigma^2\big(g(s)\big)\,L'\big(g(s)\big)}{2\,c\big(g(s)\big)\Big(L(s)-L\big(g(s)\big)\Big)} .$$

Thus, to each solution $s\mapsto g(s)$ of the equation (2.3.21) corresponds a function $(x,s)\mapsto V_g(x,s)$ defined by (2.3.20) which solves the system (2.3.8)-(2.3.11) in the region $g(s)<x<s$, and coincides with the expectation in (2.3.13) whenever $E_{x,s}(S_{\tau_g})$ and $E_{x,s}(\int_0^{\tau_g} c(X_t)\ dt)$ are finite

(the latter is easily verified by Itô formula). We shall use this fact in the proof of Theorem 2.3.1 below upon approximating the selected solution of (2.3.21) by solutions which hit the diagonal in $I\!R^2$.

6. Observe that among all possible functions $s \mapsto g(s)$, only those which satisfy (2.3.21) have the smooth-fit property (2.3.18) for $V_g(x,s)$ of (2.3.13), and vice versa. Thus *the differential equation (2.3.21) is obtained by the principle of smooth-fit* in the problem (2.2.4). The fundamental question to be answered is how to chose the optimal stopping boundary $s \mapsto g_*(s)$ among all admissible candidates which solve (2.3.21).

Before passing to answer this question let us also observe from (2.3.20) that:

$$\frac{\partial V_g}{\partial x}(x,s) = L'(x) \int_{g(s)}^{x} c(y)\ m(dy) \tag{2.3.22}$$

$$V_g'(s) = L'(s) \int_{g(s)}^{s} c(y)\ m(dy)\ . \tag{2.3.23}$$

These equations show that, in addition to the continuity of the derivative of $V_g(x,s)$ along the vertical line across $g(s)$ in (2.3.18), we have obtained the continuity of $V_g(x,s)$ along the vertical line and the diagonal in $I\!R^2$ across the point where they meet. In fact, we see that the latter condition is equivalent to the former, and thus may be used as an alternative way of looking at the principle of smooth-fit in this problem.

7. In view of the analysis about (2.3.7), we assign a constant value to $V_g(x,s)$ at all $x < g(s)$. The following properties of the solution $V_g(x,s)$ obtained are then straightforward:

(2.3.24) $\quad V_g(x,s) = s$ for $x \le g(s)$

(2.3.25) $\quad x \mapsto V_g(x,s)$ is (strictly) increasing on $[g(s), s]$

(2.3.26) $\quad (x,s) \mapsto V_g(x,s)$ is C^2 outside $\{\,(g(s),s) \mid s \in I\!R\,\}$

(2.3.27) $\quad x \mapsto V_g(x,s)$ is C^1 at $g(s)$.

Let us also make the following observations:

(2.3.28) $\quad g \mapsto V_g(x,s)$ is (strictly) *decreasing*

(2.3.29) $\quad$ The function $(a,x,s) \mapsto V_g(x,s) - a$ is *superharmonic* for the Markov process $Z_t = (A_t, X_t, S_t)$ (with respect to stopping times τ satisfying (2.2.5)).

The property (2.3.28) is evident from (2.3.20), whereas (2.3.29) is derived in the proof of Theorem 2.3.1 (see (2.3.37) below).

8. Combining (2.3.6)+(2.3.28)+(2.3.29) with the superharmonic characterization of the payoff from the Markovian theory (see [65; p.124]), and recalling the result of Proposition 2.2.1, we are led to the following Markovian principle for determining the optimal stopping boundary. We say that $s \mapsto g_*(s)$ is *an optimal stopping boundary* for the problem (2.2.4), if the stopping time τ_*

defined in (2.3.7) is optimal for this problem.

The Maximality Principle:
The optimal stopping boundary $s \mapsto g_*(s)$ *for the problem (2.2.4) is the maximal solution of the differential equation (2.3.21) which stays strictly below the diagonal in* $I\!R^2$ *(see Figure 2.1 below).*

This principle is equivalent to the superharmonic characterization of the payoff (for the process $Z_t = (A_t, X_t, S_t)$), and may be viewed as its alternative (analytic) description. The proof of its validity is given in the next theorem, the main result of the chapter.

Theorem 2.3.1 (***Optimal stopping of the maximum process***)
In the setting of (2.2.1)-(2.2.3) consider the optimal stopping problem (2.2.4) where the supremum is taken over all stopping times τ *of* X *satisfying (2.2.5).*

(I): Let $s \mapsto g_*(s)$ *denote the maximal solution of (2.3.21) which stays strictly below the diagonal in* $I\!R^2$ *(whenever such a solution exists; see Figure 2.1 below). Then we have:*

1. The payoff is finite and is given by:

$$V_*(x,s) = s + \int_{g_*(s)}^{x} \big(L(x) - L(y)\big) c(y)\ m(dy) \tag{2.3.30}$$

for $g_*(s) \le x \le s$ *, and* $V_*(x,s) = s$ *for* $x \le g_*(s)$ *.*

2. The stopping time:

$$\tau_* = \inf \{\, t > 0 \mid X_t \le g_*(S_t) \,\} \tag{2.3.31}$$

is optimal for the problem (2.2.4) whenever it satisfies (2.2.5); otherwise it is "approximately" optimal in the sense described in the proof below.

3. If there exists an optimal stopping time σ *in (2.2.4) satisfying (2.2.5), then* $P_{x,s}(\tau_* \le \sigma) = 1$ *for all* (x,s) *, and* τ_* *is an optimal stopping time for (2.2.4) as well.*

(II): If there is no (maximal) solution of (2.3.21) which stays strictly below the diagonal in $I\!R^2$ *, then* $V_*(x,s) = +\infty$ *for all* (x,s) *, and there is no optimal stopping time.*

Proof. (I): Let $s \mapsto g(s)$ be any solution of (2.3.21) satisfying $g(s) < s$ for all s . Then, as indicated above, the function $V_g(x,s)$ defined by (2.3.20) solves the system (2.3.8)-(2.3.11) in the region $g(s) < x < s$. Due to (2.3.26) and (2.3.27), Itô formula can be applied to the process $V_g(X_t, S_t)$, and in this way by (2.2.1)-(2.2.2) we get:

$$\begin{aligned} V_g(X_t,S_t) &= V_g(x,s) + \int_0^t \frac{\partial V_g}{\partial x}(X_r,S_r)\ dX_r \\ &\quad + \int_0^t \frac{\partial V_g}{\partial s}(X_r,S_r)\ dS_r + \frac{1}{2}\int_0^t \frac{\partial^2 V_g}{\partial x^2}(X_r,S_r)\ d\big\langle X,X\big\rangle_r \\ &= V_g(x,s) + \int_0^t \sigma(X_r)\,\frac{\partial V_g}{\partial x}(X_r,S_r)\ dB_r + \int_0^t (I\!L_X V_g)(X_r,S_r)\ dr \end{aligned} \tag{2.3.32}$$

where the integral with respect to dS_r is zero, since the increment ΔS_r outside the diagonal in

$I\!R^2$ equals zero, while at the diagonal we have (2.3.9).

The process $M = (M_t)_{t \geq 0}$ defined by:

$$M_t = \int_0^t \sigma(X_r) \frac{\partial V_g}{\partial x}(X_r, S_r) \; dB_r \tag{2.3.33}$$

is a continuous local martingale. Introducing the increasing process:

$$P_t = \int_0^t c(X_r) \, 1_{\left(X_r \leq g(S_r)\right)} \; dr \tag{2.3.34}$$

and using the fact that the set of all t for which X_t is either $g(S_t)$ or S_t is of Lebesgue measure zero, the identity (2.3.32) can be rewritten as:

$$V_g(X_t, S_t) - \int_0^t c(X_r) \; dr = V_g(x,s) + M_t - P_t \tag{2.3.35}$$

by means of (2.3.8) with (2.3.24). From this representation we see that the process $V_g(X_t, S_t) - \int_0^t c(X_r) \; dr$ is a local supermartingale.

Let τ be any stopping time of X satisfying (2.2.5). Choose a localization sequence $(\sigma_n)_{n \geq 1}$ of bounded stopping times for M. By means of (2.3.24) and (2.3.25) we see that $V_g(x,s) \geq s$ for all (x,s), so that from (2.3.35) it follows:

$$\begin{aligned} E_{x,s}\left(S_{\tau\wedge\sigma_n} - \int_0^{\tau\wedge\sigma_n} c(X_t) \; dt\right) &\leq E_{x,s}\left(V_g(X_{\tau\wedge\sigma_n}, S_{\tau\wedge\sigma_n}) - \int_0^{\tau\wedge\sigma_n} c(X_t) \; dt\right) \\ &\leq V_g(x,s) + E_{x,s}(M_{\tau\wedge\sigma_n}) = V_g(x,s) \; . \end{aligned} \tag{2.3.36}$$

Letting $n \rightarrow \infty$, and using Fatou's lemma with (2.2.5), we get:

$$E_{x,s}\left(S_\tau - \int_0^\tau c(X_t) \; dt\right) \leq V_g(x,s) \; . \tag{2.3.37}$$

This proves (2.3.29). Taking the supremum over all such τ, and then the infimum over all such g, by means of (2.3.28) we may conclude:

$$V_*(x,s) \leq \inf_g \; V_g(x,s) = V_{g_*}(x,s) \tag{2.3.38}$$

for all (x,s). From these considerations it clearly follows that the only possible candidate for the optimal stopping boundary is the maximal solution $s \mapsto g_*(s)$ of (2.3.21).

To prove that we have the equality in (2.3.38), and that the payoff $V_*(x,s)$ is given by (2.3.30), assume first that the stopping time τ_* defined by (2.3.31) satisfies (2.2.5). Then, as pointed out when deriving (2.3.20), we have:

$$V_{g_*}(x,s) = E_{x,s}\left(S_{\tau_{g_*}} - \int_0^{\tau_{g_*}} c(X_t) \; dt\right) \tag{2.3.39}$$

so that $V_{g_*}(x,s) = V_*(x,s)$ in (2.3.38) and τ_* is an optimal stopping time. The explicit expression given in (2.3.30) is obtained by (2.3.20).

Assume now that τ_* fails to satisfy (2.2.5) . Let $(g_n)_{n\geq 1}$ be a decreasing sequence of solutions of (2.3.21) satisfying $g_n(s) \downarrow g_*(s)$ as $n \to \infty$ for all s . Note that each such solution must hit the diagonal in $I\!R^2$, so the stopping times τ_{g_n} defined as in (2.3.12) must satisfy (2.2.5). Moreover, since $S_{\tau_{g_n}}$ is bounded by a constant, we see that $V_{g_n}(x,s)$ defined as in (2.3.13) is given by (2.3.20) with $g = g_n$ for $n \geq 1$. By letting $n \to \infty$ we get:

$$V_{g_*}(x,s) = \lim_{n\to\infty} V_{g_n}(x,s) = \lim_{n\to\infty} E_{x,s}\left(S_{\tau_{g_n}} - \int_0^{\tau_{g_n}} c(X_t)\ dt \right) . \tag{2.3.40}$$

This shows that the equality in (2.3.38) is attained through the sequence of stopping times $(\tau_{g_n})_{n\geq 1}$, and the explicit expression in (2.3.30) is easily obtained as already indicated above.

To prove the final (uniqueness) statement, assume that σ is an optimal stopping time in (2.2.4) satisfying (2.2.5). Suppose that $P_{x,s}(\sigma < \tau_*) > 0$. Note that τ_* can be written in the form:

$$\tau_* = \inf\{\, t > 0 \mid V_*(X_t, S_t) = S_t \,\} \tag{2.3.41}$$

so that $S_\sigma < V_*(X_\sigma, S_\sigma)$ on $\{\sigma < \tau_*\}$, and thus:

$$E_{x,s}\left(S_\sigma - \int_0^\sigma c(X_t)\ dt \right) < E_{x,s}\left(V_*(X_\sigma, S_\sigma) - \int_0^\sigma c(X_t)\ dt \right) \leq V_*(x,s) \tag{2.3.42}$$

where the latter inequality is derived as in (2.3.37), since the process $V_*(X_t,S_t) - \int_0^t c(X_r)\ dr$ is a local supermartingale. The strict inequality in (2.3.42) shows that $P_{x,s}(\sigma < \tau_*) > 0$ fails, so we must have $P_{x,s}(\tau_* \leq \sigma) = 1$ for all (x,s) .

To prove the optimality of τ_* in such a case, it is enough to note that if σ satisfies (2.2.5) then τ_* must satisfy it as well. Therefore (2.3.39) is satisfied, and thus τ_* is optimal. A straightforward argument can also be given by using the local supermartingale property of the process $V_*(X_t,S_t) - \int_0^t c(X_r)\ dr$; since $P_{x,s}(\tau_* \leq \sigma) = 1$, we get:

$$\begin{aligned} V_*(x,s) &= E_{x,s}\left(S_\sigma - \int_0^\sigma c(X_t)\ dt \right) \leq E_{x,s}\left(V_*(X_\sigma, S_\sigma) - \int_0^\sigma c(X_t)\ dt \right) \\ &\leq E_{x,s}\left(V_*(X_{\tau_*}, S_{\tau_*}) - \int_0^{\tau_*} c(X_t)\ dt \right) = E_{x,s}\left(S_{\tau_*} - \int_0^{\tau_*} c(X_t)\ dt \right) \end{aligned} \tag{2.3.43}$$

so τ_* is optimal for (2.2.4). The proof of the first part of the theorem is complete.

(II): Let $(g_n)_{n\geq 1}$ be a decreasing sequence of solutions of (2.3.21) which satisfy $g_n(0) = -n$ for $n \geq 1$. Then each g_n must hit the diagonal in $I\!R^2$ at some $s_n > 0$ for which we have $s_n \uparrow \infty$ when $n \to \infty$. Since there is no solution of (2.3.21) which stays below the diagonal, we must have $g_n(s) \downarrow -\infty$ as $n \to \infty$ for all s . Let τ_{g_n} denote the stopping time defined by (2.3.12) with $g = g_n$. Then τ_{g_n} satisfies (2.2.5), and since $S_{\tau_{g_n}} \leq s \vee s_n$, we see that $V_{g_n}(x,s)$ defined by (2.3.13) with $g = g_n$ is given as in (2.3.20):

$$V_{g_n}(x,s) = s + \int_{g_n(s)}^x \big(L(x) - L(y)\big) c(y)\ m(dy) \tag{2.3.44}$$

for all $g_n(s) \le x \le s$. Letting $n \to \infty$ in (2.3.44), we see that the following integral:

$$I := \int_{-\infty}^{x} \big(L(x) - L(y)\big) c(y)\ m(dy) \tag{2.3.45}$$

plays a crucial role in the proof (independently of the given x and s).

Assume first that $I = +\infty$ (this is the case whenever $c(y) \ge \varepsilon > 0$ for all y , and $-\infty$ is a natural boundary point for X). Then from (2.3.44) we clearly get:

$$V_*(x,s) \ge \lim_{n \to \infty} V_{g_n}(x,s) = +\infty \tag{2.3.46}$$

so the payoff must be infinite.

On the other hand, if $I < \infty$, then (2.2.11)+(2.2.12) imply:

$$E_{x,s}\bigg(\int_0^{\tau_{\hat{s}}} c(X_t)\ dt \bigg) \le \int_{-\infty}^{\hat{s}} \big(L(\hat{s}) - L(y)\big) c(y)\ m(dy) < \infty \tag{2.3.47}$$

where $\tau_{\hat{s}} = \inf\{\, t > 0 \mid X_t = \hat{s} \,\}$ for $\hat{s} \ge s$. Thus, if we let the process (X_t, S_t) first hit $(\hat{s}, \hat{s})$, and then the boundary $\{\, (g_n(s), s) \mid s \in I\!R \,\}$ with $n \to \infty$, then by (2.3.44) (with $x = s = \hat{s}$) we see that the payoff equals at least $\hat{s}$. More precisely, if the process (X_t, S_t) starts at (x, s) , consider the stopping times $\tau_n = \tau_{\hat{s}} + \tau_{g_n} \circ \theta_{\tau_{\hat{s}}}$ for $n \ge 1$. Then by (2.3.47) we see that each τ_n satisfies (2.2.5), and by the strong Markov property of X we easily get:

$$V_*(x,s) \ge \limsup_{n \to \infty}\ E_{x,s}\bigg(S_{\tau_n} - \int_0^{\tau_n} c(X_t)\, dt \bigg) \ge \hat{s}\ . \tag{2.3.48}$$

By letting $\hat{s} \uparrow \infty$, we again find $V_*(x,s) = +\infty$. The proof of the theorem is complete. □

9. ***On the equation (2.3.21).*** Theorem 2.3.1 shows that the optimal stopping problem (2.2.4) reduces to the problem of solving the first-order nonlinear differential equation (2.3.21). If this equation admits a maximal solution which stays strictly below the diagonal in $I\!R^2$, then this solution is an optimal stopping boundary.

We may note that this equation is of the following *normal* form:

$$y' = \frac{F(y)}{G(x) - G(y)} \tag{2.3.49}$$

for $x > y$, where $y \mapsto F(y)$ is strictly positive, and $x \mapsto G(x)$ is strictly increasing. To the best of our knowledge the equation (2.3.49) has not been studied before, and in view of the result proved above we want to point out the need for its investigation. It turns out that its treatment depends heavily on the behaviour of the map G .

(i): If the process X is in natural scale, that is $L(x) = x$ for all x , we can completely characterize and describe the maximal solution of (2.3.21). This can be done in terms of the equation (2.3.49) with $G(x) = x$ and $F(y) = \sigma^2(y)/2c(y)$ as follows. Note that by passing to the inverse $z \mapsto y^{-1}(z)$, the equation (2.3.49) in this case can be rewritten as:

$$(2.3.50) \qquad \big(y^{-1}\big)'(z) - \frac{1}{F(z)}\, y^{-1}(z) = -\frac{z}{F(z)}\ .$$

This is a first-order linear equation and its general solution is given by:

$$(2.3.51) \qquad y_\alpha^{-1}(z) = \exp\left(\int_0^z \frac{dy}{F(y)}\right)\left(\alpha - \int_0^z \frac{y}{F(y)}\, \exp\left(-\int_0^y \frac{du}{F(u)}\right) dy\right)$$

where α is a constant. Hence, with $G(x) = x$, *the necessary and sufficient condition* for the equation (2.3.49) to admit a maximal solution which stays strictly below the diagonal in $I\!R^2$, is that:

$$(2.3.52) \qquad \alpha_* := \sup_{z \in I\!R}\left(z\exp\left(-\int_0^z \frac{dy}{F(y)}\right) + \int_0^z \frac{y}{F(y)}\, \exp\left(-\int_0^y \frac{du}{F(u)}\right) dy\right) < \infty$$

and that this supremum is not attained at any $z \in I\!R$. In this case the maximal solution $x \mapsto y_*(x)$ of (2.3.49) can be expressed explicitly through its inverse $z \mapsto y_{\alpha_*}^{-1}(z)$ given by (2.3.51).

Note also when $L(x) = G(x) = x^2 sign(x)$ that the same argument transforms (2.3.49) into a *Riccati equation*, which then can be further transformed into a linear homogeneous equation of second order by means of standard techniques. The trick of passing to the inverse in (2.3.21) is further used in [P9] where a natural connection between the result of the present chapter and the Azéma-Yor solution of the Skorokhod-embedding problem [2] is described.

(ii): If the process X is not in natural scale, then the treatment of (2.3.49) is much harder, due to the lack of closed form solutions. In such cases it is possible to prove (or disprove) the existence of the maximal solution by using Picard's method of successive approximations. The idea is to use Picard's theorem locally, step-by-step, and in this way show the existence of some global solution which stays strictly below the diagonal. Then, by passing to the equivalent integral equation and using a monotone convergence theorem, one can argue that this implies the existence of the maximal solution. This technique is described in detail in Section 3 of [P11] in the case of $G(x) = x^p$ and $F(y) = y^{p+1}$ when $p > 1$. It is also seen there that during the construction one obtains tight bounds on the maximal solution which makes it possible to compute it numerically as accurate as desired (see [P11] for details). In this process it is desirable to have a local existence and uniqueness of the solution, and these are provided by the following general facts.

From the general theory (Picard's method) we know that if the direction field $(x,y) \mapsto f(x,y) := F(y)/(G(x) - G(y))$ is (locally) continuous and (locally) Lipschitz in the second variable, then the equation (2.3.49) admits (locally) a unique solution. For instance, this will be so if along a (local) continuity of $(x,y) \mapsto f(x,y)$, we have a (local) continuity of $(x,y) \mapsto (\partial f/\partial y)(x,y)$. In particular, upon differentiating over y in $f(x,y)$ we see that (2.3.21) admits (locally) a unique solution whenever the map $y \mapsto \sigma^2(y)L'(y)/c(y)$ is (locally) C^1 . It is also possible to prove that the equation (2.3.49) admits (locally) a solution, if only the (local) continuity of the direction field $(x,y) \mapsto F(y)/(G(x) - G(y))$ is verified. However, such a solution may fail to be (locally) unique.

Instead of entering further into such abstract considerations here, we shall rather confine ourselves to some concrete examples with applications in the next section.

10. We have proved in Theorem 2.3.1 that τ_* is optimal for (2.2.4) whenever it satisfies (2.2.5). In Example 2.4.1 below we will exhibit a stopping time τ_* which fails to satisfy (2.2.5), but nevertheless its payoff is given by (2.3.30) as proved above. In this case τ_* is "approximately"

optimal in the sense that (2.3.40) holds with $\tau_{g_n} \uparrow \tau_*$ as $n \to \infty$.

11. ***Other state spaces.*** The result of Theorem 2.3.1 extends to diffusions with other state spaces in $I\!R$. In view of many applications, we will indicate such an extension for non-negative diffusions.

In the setting of (2.2.1)-(2.2.3) assume that the diffusion X is non-negative, consider the optimal stopping problem (2.2.4) where the supremum is taken over all stopping times τ of X satisfying (2.2.5), and note that the result of Proposition 2.2.1 extends to this case provided that the diagonal is taken in $]0,\infty[^2$. In this context it is natural to assume that $\sigma(x) > 0$ for $x > 0$, and $\sigma(0)$ may be equal 0 . Similarly, we shall see that the case of strictly positive cost function c differs from the case when c is strictly positive only on $]0,\infty[$. In any case, both $x \mapsto \sigma(x)$ and $x \mapsto c(x)$ are assumed continuous on $[0,\infty[$.

In addition to the infinitesimal characteristics from (2.2.1) which govern X in $]0,\infty[$, we must specify the boundary behaviour of X at 0 . For this we shall consider the cases when 0 is a natural, exit, regular (instantaneously reflecting), and entrance boundary point (see [46; pp.226-250]).

The relevant fact in the case when 0 is either a *natural* or *exit* boundary point is that:

$$\int_0^s \big(L(s)-L(y)\big)\, c(y)\ m(dy) = +\infty \tag{2.3.53}$$

for all $s > 0$ whenever $c(0) > 0$. In view of (2.3.30) this shows that for the maximal solution of (2.3.21) we must have $0 < g_*(s) < s$ for all $s > 0$ unless $V_*(s,s) = +\infty$. If $c(0) = 0$, then the integral in (2.3.53) can be finite, and we cannot state a similar claim; but from our method used below it will be clear how to handle such a case too, and therefore the details in this direction will be omitted for simplicity.

The relevant fact in the case when 0 is either a *regular (instantaneously reflecting)* or *entrance* boundary point is that:

$$E_{0,s}\bigg(\int_0^{\tau_{s_*}} c(X_t)\ dt\bigg) = \int_0^{s_*} \big(L(s_*)-L(y)\big)\, c(y)\ m(dy) \tag{2.3.54}$$

for all $s_* \ge s > 0$ where $\tau_{s_*} = \inf\{\, t > 0 \mid X_t = s_* \,\}$. In view of (2.3.30) this shows that it is never optimal to stop at $(0,s)$. Therefore, if the maximal solution of (2.3.21) satisfies $g_*(s_*) = 0$ for some $s_* > 0$ with $g_*(s) > 0$ for all $s > s_*$, then $\tau_* = \inf\{\, t > 0 \mid X_t \le g_*(S_t) \,\}$ is to be the optimal stopping time, since X does not take negative values. If moreover $c(0) = 0$, then the value of $m(\{0\})$ does not play any role, and all regular behaviour (from absorption $m(\{0\}) = +\infty$, over sticky barrier phenomenon $0 < m(\{0\}) < +\infty$, to instantaneous reflection $m(\{0\}) = 0$) can be treated in the same way.

For simplicity in the next result we will assume that $c(0) > 0$ if 0 is either a natural (attracting or unattainable) or an exit boundary point, and will only consider the instantaneously-reflecting regular case. The remaining cases can be treated similarly.

Corollary 2.3.2 (***Optimal stopping for non-negative diffusions***)

In the setting of (2.2.1)-(2.2.3) assume that the diffusion X is non-negative, and that 0 is a natural, exit, instantaneously-reflecting regular, or entrance boundary point. Consider the optimal stopping problem (2.2.4) where the supremum is taken over all stopping times τ of X satisfying

(2.2.5).

(I): Let $s \mapsto g_*(s)$ *denote the maximal solution of (2.3.21) in the following sense (whenever such a solution exists; see Figure 2.2 below): There exists a point* $s_* \geq 0$ *(with* $s_* = 0$ *if* 0 *is either a natural or an exit boundary point) such that* $g_*(s_*) = 0$ *and* $g_*(s) > 0$ *for all* $s > s_*$ *; the map* $s \mapsto g_*(s)$ *solves (2.3.21) for* $s > s_*$ *and stays strictly below the diagonal in* $]0, \infty[^2$ *; the map* $s \mapsto g_*(s)$ *is the maximal solution satisfying these two properties (the comparison of two maps is taken pointwise wherever they are both strictly positive). Then we have:*

1. The payoff is finite and for $s \geq s_*$ *is given by:*

$$V_*(x,s) = s + \int_{g_*(s)}^{x} \big(L(x)-L(y)\big)c(y)\ m(dy) \tag{2.3.55}$$

for $g_*(s) \leq x \leq s$ *with* $V_*(x,s) = s$ *for* $0 \leq x \leq g_*(s)$ *, and for* $s \leq s_*$ *(when* 0 *is either an instantaneously-reflecting regular or an entrance boundary point) is given by:*

$$V_*(x,s) = s_* + \int_0^x \big(L(x)-L(y)\big)c(y)\ m(dy) \tag{2.3.56}$$

for $0 \leq x \leq s$ *.*

2. The stopping time:

$$\tau_* = \inf\big\{\, t > 0 \mid S_t \geq s_*\,,\ X_t \leq g_*(S_t) \,\big\} \tag{2.3.57}$$

is optimal for the problem (2.2.4) whenever it satisfies (2.2.5); otherwise, it is "approximately" optimal.

3. If there exists an optimal stopping time σ *in (2.2.4) satisfying (2.2.5), then* $P_{x,s}(\tau_* \leq \sigma) = 1$ *for all* (x,s) *, and* τ_* *is an optimal stopping time for (2.2.4) as well.*

(II): If there is no (maximal) solution of (2.3.21) in the sense of (I) above, then $V_*(x,s) = +\infty$ *for all* (x,s) *, and there is no optimal stopping time.*

Proof. Only with minor changes the proof can be carried out in exactly the same way as the proof of Theorem 2.3.1 upon using the additional facts about (2.3.53) and (2.3.54) stated above, and the details will be omitted; note, however, that in the case when 0 is either an instantaneously-reflecting regular or an entrance boundary point, the strong Markov property of X at $\tau_{s_*} = \inf\{\, t > 0 \mid X_t = s_* \,\}$ gives:

$$V_*(x,s) = s_* + \int_0^{s_*} \big(L(s_*)-L(y)\big)c(y)\ m(dy) - E_{x,s}\bigg(\int_0^{\tau_{s_*}} c(X_t)\ dt\bigg) \tag{2.3.58}$$

for all $0 \leq x \leq s \leq s_*$. Hence formula (2.3.56) follows by applying (2.2.11)+(2.2.12) to the last term in (2.3.58). (In the instantaneous reflecting case one can make use of τ_{s_*,s_*} after extending L to $I\!R_-$ by setting $L(x) := -L(-x)$ for $x < 0$). The proof is complete. □

12. ***The "discounted" problem.*** One is often more interested in the discounted version of the optimal stopping problem (2.2.4). Such a problem can be reduced to the initial problem (2.2.4) by changing the underlying diffusion process.

Given a continuous function $x \mapsto \lambda(x) \geq 0$ called the *discounting rate*, in the setting of

(2.2.1)-(2.2.3) introduce the functional:

$$\Lambda(t) = \int_0^t \lambda(X_r)\ dr\ , \tag{2.3.59}$$

and consider the optimal stopping problem with payoff:

$$V_*(x,s) \ = \ \sup_\tau \ E_{x,s}\Big(e^{-\Lambda(\tau)} S_\tau - \int_0^\tau e^{-\Lambda(t)} c(X_t)\, dt \Big) \tag{2.3.60}$$

where the supremum is taken over all stopping times τ of X for which the integral has finite expectation, and the *cost* function $x \mapsto c(x) > 0$ is continuous.

The standard argument shows that the problem (2.3.60) is equivalent to the problem:

$$V_*(x,s) \ = \ \sup_\tau \ E_{x,s}\Big(\widetilde{S}_\tau - \int_0^\tau c(\widetilde{X}_t)\, dt \Big) \tag{2.3.61}$$

where $\widetilde{X} = (\widetilde{X}_t)_{t\ge 0}$ is a diffusion process which corresponds to the "killing" of the sample paths of X at the "rate" $\lambda(X)$. The infinitesimal operator of $\widetilde{X}$ is given by:

$$I\!\!L_{\widetilde{X}} = -\lambda(x) + \mu(x)\,\frac{\partial}{\partial x} + \frac{\sigma^2(x)}{2}\,\frac{\partial^2}{\partial x^2}\ . \tag{2.3.62}$$

We conjecture that the maximality principle proved above also holds for this problem (see [60]). The main technical difficulty in a general treatment of this problem is the fact that the infinitesimal operator $I\!\!L_{\widetilde{X}}$ has the constant term $-\lambda(x)$, so that $I\!\!L_{\widetilde{X}} = 0$ may have no simple solution. Nonetheless, it is clear that the corresponding system (2.3.8)-(2.3.11) must be valid, and this system defines the (maximal) boundary $s \mapsto g_*(s)$ implicitly.

13. ***The "Markovian" cost problem***. Yet another class of optimal stopping problems reduces to the problem (2.2.4). Suppose that in the setting of (2.2.1)-(2.2.3) we are given a smooth function $x \mapsto D(x)$, and consider the optimal stopping problem with payoff:

$$V_*(x,s) \ = \ \sup_\tau \ E_{x,s}\Big(S_\tau - D(X_\tau) \Big) \tag{2.3.63}$$

where the supremum is taken over a class of stopping times τ of X . Then a variant of Itô formula applied to $D(X_t)$, the optional sampling theorem applied to the continuous local martingale $M_t = \int_0^t D'(X_s)\,\sigma(X_s)\,dB_s$ localized if necessary, and uniform integrability conditions enable one to conclude:

$$E_{x,s}\Big(D(X_\tau) \Big) = D(x) + E_{x,s}\Big(\int_0^\tau \big(I\!\!L_X D\big)(X_s)\, ds \Big)\ . \tag{2.3.64}$$

Hence we see that the problem (2.3.63) reduces to the problem (2.2.4) with $x \mapsto c(x)$ replaced by $x \mapsto (I\!\!L_X D)(x)$ whenever non-negative. The conditions assumed above to make such a transfer possible are not restrictive in general (see Example 2.4.2 below).

2.4 Examples and applications

There is a large number of applications of the optimal stopping results (Theorem 2.3.1 and Corollary 2.3.2) from the previous section. In this section we present some of them (see also

[P9]). Our main aim is to derive sharp versions of some known classical inequalities, as well as to deduce some new sharp inequalities closely related. It should be noted that the method applies to all diffusions. Throughout $B=(B_t)_{t\ge 0}$ denotes the standard Brownian motion started at zero.

Example 2.4.1 (***The Doob inequality***) Consider the optimal stopping problem (2.2.4) with $X_t=|B_t+x|^p$ and $c(x)=c\,x^{(p-2)/p}$ for $p>1$. Then X is a non-negative diffusion having 0 as an instantaneously-reflecting regular boundary point, and the infinitesimal operator of X in $]0,\infty[$ is given by the expression:

$$\mathbb{L}_X = \frac{p(p-1)}{2}\,x^{1-2/p}\,\frac{\partial}{\partial x} + \frac{p^2}{2}\,x^{2-2/p}\,\frac{\partial^2}{\partial x^2}\ . \tag{2.4.1}$$

The equation (2.3.21) takes the form:

$$g'(s) = \frac{p g^{1/p}(s)}{2c\Big(s^{1/p}-g^{1/p}(s)\Big)} \tag{2.4.2}$$

and its maximal solution of (2.4.2) is given by:

$$g_*(s) = \alpha\, s \tag{2.4.3}$$

where $0<\alpha<1$ is the maximal root (out of two possible ones) of the equation:

$$\alpha - \alpha^{1-1/p} + p/2c = 0\ . \tag{2.4.4}$$

It is easily verified that (2.4.4) admits such a root if and only if $c\ge p^{p+1}/2(p-1)^{(p-1)}$. Then by the result of Corollary 2.3.2, upon using (2.3.64) and letting $c\downarrow p^{p+1}/2(p-1)^{(p-1)}$, we get:

$$E\Big(\max_{0\le t\le\tau}|B_t+x\,|^p\Big) \le \Big(\frac{p}{p-1}\Big)^p E|B_\tau+x\,|^p - \Big(\frac{p}{p-1}\Big)x^p \tag{2.4.5}$$

for all stopping times τ of B such that $E(\tau^{p/2})<\infty$. The constants $(p/(p-1))^p$ and $p/(p-1)$ are the best possible, and the equality in (2.4.5) is attained in the limit through the stopping times $\tau_*=\inf\{\,t>0\mid X_t\le\alpha S_t\,\}$ when $c\downarrow p^{p+1}/2(p-1)^{(p-1)}$. These stopping times are pointwise the smallest possible with this property, and they satisfy $E(\tau_*^{p/2})<\infty$ if and only if $c>p^{p+1}/2(p-1)^{(p-1)}$. For more information and remaining details we refer to [P5].

The inequality (2.4.5) can be further extended (for simplicity we state this extension only for $x=0$) by using the result of Corollary 2.3.2 as follows:

$$E\Big(\max_{0\le t\le\tau}|B_t|^p\Big) \le \gamma^*_{p,q}\Big(E\int_0^\tau |B_t|^{q-1}\,dt\Big)^{p/(q+1)} \tag{2.4.6}$$

for all stopping times τ of B, all $0<p<1+q$, and all $q>0$, with the best possible value for the constant $\gamma^*_{p,q}$ being equal:

$$\gamma^*_{p,q} = (1+\kappa)\Big(\frac{s_*}{\kappa^\kappa}\Big)^{1/(1+\kappa)} \tag{2.4.7}$$

where $\kappa=p/(q-p+1)$, and s_* is the zero point of the maximal solution $s\mapsto g_*(s)$ of:

$$(2.4.8) \qquad g'(s) = \frac{p g^{(1-q/p)}(s)}{2\Big(s^{1/p} - g^{1/p}(s)\Big)}$$

satisfying $0 < g_*(s) < s$ for all $s > s_*$. (This solution is also characterized by $g_*(s)/s \to 1$ for $s \to \infty$.) The equality in (2.4.6) is attained at the stopping time $\tau_* = \inf\{\, t > 0 \mid X_t = g_*(S_t)\,\}$ which is pointwise the smallest possible with this property. In the case $p = 1$ the closed form for $s \mapsto g_*(s)$ is found:

$$(2.4.9) \qquad s \exp\left(-\frac{2}{pq}\, g_*^q(s)\right) + \frac{2}{p}\int_0^{g_*(s)} t^q \exp\left(-\frac{2}{pq}\, t^q\right)\, dt = \left(\frac{pq}{2}\right)^{1/q} \Gamma\left(\frac{q+1}{q}\right)$$

for $s \geq s_*$. This, in particular, yields:

$$(2.4.10) \qquad \gamma_{1,q}^* = \left(\frac{q(1+q)}{2}\right)^{1/(1+q)} \left(\Gamma\left(2 + \frac{1}{q}\right)\right)^{q/(1+q)}$$

for all $q > 0$. In the case $p \neq 1$ no closed form for $s \mapsto g_*(s)$ seems to exist. For more information and remaining details in this direction, as well as for the extension of inequality (2.4.6) to $x \neq 0$, we refer to [P6] (see also [P14]). To give a more familiar form to the inequality (2.4.6), note by Itô formula and the optional sampling theorem that:

$$(2.4.11) \qquad E\left(\int_0^\tau |B_t|^{q-1} dt\right) = \frac{2}{q(q+1)} E|B_\tau|^{q+1}$$

whenever τ is a stopping time of B satisfying $E(\tau^{(q+1)/2}) < \infty$ for $q > 0$. Hence we see that the right-hand side in (2.4.6) is the well-known Doob's bound. The advantage of formulation (2.4.6) lies in its validity for all stopping times.

While the inequality (2.4.6) (with some constant $\gamma_{p,q} > 0$) can be derived quite easily, the question of its sharpness has gained interest. The case $p = 1$ was treated independently by Jacka [40] (probabilistic methods) and Gilat [33] (analytic methods) who both found the best possible value $\gamma_{1,q}^*$ for $q > 0$. This in particular yields $\gamma_{1,1}^* = \sqrt{2}$ which was independently obtained by Dubins and Schwarz [23], and later again by Dubins, Shepp and Shiryaev [24] who studied a more general case of Bessel processes. (A simple probabilistic proof for $\gamma_{1,1}^* = \sqrt{2}$ is given in [P15]). The Bessel processes results are further extended in [55]. In the case $p = 1 + q$ with $q > 0$, the inequality (2.4.6) reduces to the Doob's maximal inequality (2.4.5). I learned from D. Burkholder that this inequality can be obtained as a by-product from his new proof of Doob's inequality for discrete non-negative submartingales (see [15; p.14]). The proof given there in essence relies on a submartingale property, while the proof presented above in essence relies on a strong Markov property. Cox [20] also derived the analogue of this inequality for discrete martingales by a method which is based on results from the theory of moments. That the equality in Doob's maximal inequality (2.4.5) cannot be attained by a non-zero (sub)martingale was observed by Cox [20]. It should be noted that this fact also follows from the method and results above (the equality in (2.4.5) is attained only in the limit). The best values $\gamma_{p,q}^*$ in (2.4.6) and the corresponding optimal stopping times τ^* for all $0 < p \leq 1 + q$ and all $q > 0$ are given in [P6]. The main novelty about (2.4.5) and (2.4.6) which is realised here is that the optimal τ_* from (2.3.57) is pointwise the smallest possible stopping time at which the equalities in (2.4.5) (in the limit) and in (2.4.6) can be attained. The results about (2.4.5) and (2.4.6) extend to all non-negative submartingales. This can be obtained by using the maximal embedding result of Jacka [39] (for details see [P5] and [P6]).

Example 2.4.2 (***The Hardy-Littlewood inequality***) Consider the "Markovian" cost problem (2.3.63) with $X_t = |B_t + x|$ and $D(x) = x \log x$ for $x \geq 0$. Then X is a non-negative diffusion having 0 as an instantaneously-reflecting regular boundary point, and the infinitesimal operator of X in $]0, \infty[$ is given by (2.4.1) with $p = 1$. The main difficulty in this problem is that we cannot apply Itô formula directly to $D(X_t)$ as suggested in 13 of Section 2.3. Thus we truncate $D(x)$ by setting $\widetilde{D}(x) = D(x)$ for $x \geq 1/e$ and $\widetilde{D}(x) = -1/e$ for $0 \leq x \leq 1/e$. Then $\widetilde{D} \in C^1$ and $\widetilde{D}''$ exists and is continuous everywhere but at $1/e$. Thus the Itô-Tanaka formula can be applied, and since the time spent by X at $1/e$ is of Lebesgue measure zero, this formula reduces to Itô formula. In this way the problem (2.3.63) reduces to the problem (2.2.4). The equation (2.3.21) takes the form:

$$g'(s) = \frac{g(s)}{2c\Big(s - g(s)\Big)} \tag{2.4.12}$$

and its maximal solution of (2.4.12) is given by:

$$g_*(s) = \alpha s \tag{2.4.13}$$

where $\alpha = (c-1)/c$. By applying the result of Corollary 2.2.2 we get:

$$E\Big(\max_{0 \leq t \leq \tau} |B_t + x|\Big) \leq V(x;c) + cE\Big(|B_\tau + x| \log |B_\tau + x|\Big) \tag{2.4.14}$$

for all $c > 1$ and all stopping times τ of B satisfying $E(\tau^r) < \infty$ for some $r > 1/2$, where:

$$\begin{aligned} V(x;c) &= \frac{c^2}{e(c-1)} \quad \text{if } 0 \leq x \leq u_* \\ &= cx \log\left(\frac{c}{x(c-1)}\right) \quad \text{if } x \geq u_* \end{aligned} \tag{2.4.15}$$

with $u_* = c/e(c-1)$. This inequality is sharp, and for each $c > 1$ and $x \geq 0$ given and fixed, the equality in (2.4.14) is attained at the stopping time:

$$\tau_* = \inf \big\{\, t > 0 \mid S_t \geq u_* \,,\ X_t = \alpha S_t \,\big\} \tag{2.4.16}$$

which is pointwise the smallest possible with this property.

The same problem with more familiar $D(x) = x \log^+ x$ brings the local time of X at 1 into the consideration (see [P7] for details), and the analogue of (2.4.14) may be stated as follows:

$$E\Big(\max_{0 \leq t \leq \tau} |B_t + x|\Big) \leq V_+(x;c) + cE\Big(|B_\tau + x| \log^+ |B_\tau + x|\Big) \tag{2.4.17}$$

for all $c > 1$ and all stopping times τ of B satisfying $E(\tau^r) < \infty$ for some $r > 1/2$, where:

$$
\begin{aligned}
\text{(2.4.18)} \qquad V_{+}(x;c) \;&=\; 1+\frac{1}{e^{c}(c\!-\!1)} \quad \text{if} \;\; 0 \le x \le v_* \\
&=\; x+(1\!-\!x)\log(x\!-\!1)-(c+\log(c\!-\!1))(x\!-\!1) \quad \text{if} \;\; v_* \le x \le z_* \\
&=\; cx\log\left(\frac{c}{x(c\!-\!1)}\right) \quad \text{if} \;\; x \ge z_*
\end{aligned}
$$

with $v_* = 1+1/e^{c}(c\!-\!1)$ and $z_* = c/(c\!-\!1)$. This inequality is sharp, and for each $c > 1$ and $x \ge 0$ given and fixed, the equality in (2.4.17) is attained at the stopping time:

$$
\text{(2.4.19)} \qquad \sigma_* = \inf\; \big\{\, t>0 \;\mid\; S_t \,\ge\, v_* \;,\;\; X_t \,=\, 1 \vee \alpha S_t \,\big\}
$$

which is pointwise the smallest possible with this property. For remaining details and more information on (2.4.14) and (2.4.17) we refer to [P7]. Note that (2.4.14)-(2.4.19) contain and refine the results of Gilat [32] which settle a question raised by Dubins and Gilat [22], and later again by Cox [20], and which are obtained by analytic methods.

Example 2.4.3 (***A sharp integral inequality of the L log L-type***) Consider the optimal stopping problem (2.2.4) with $X_t = |B_t\!+\!x|$ and $c(x) = 1/(1\!+\!x)$ for $x \ge 0$. Then X is a non-negative diffusion having 0 as an instantaneously-reflecting regular boundary point, and the infinitesimal operator of X in $]0,\infty[$ is given by (2.4.1) with $p = 1$. The equation (2.3.21) takes the form:

$$
\text{(2.4.20)} \qquad g'(s) \;=\; \frac{1+g(s)}{2c\,\Big(s\!-\!g(s)\Big)}
$$

and its maximal solution of (2.4.20) is given by:

$$
\text{(2.4.21)} \qquad g_*(s) \;=\; \alpha\, s \,-\, \beta
$$

where $\alpha = (2c\!-\!1)/2c$ and $\beta = 1/2c$. By applying the result of Corollary 2.2.2 we get:

$$
\text{(2.4.22)} \qquad E\bigg(\max_{0\le t\le\tau} |B_t\!+\!x|\bigg) \;\le\; W(x;c) + cE\bigg(\int_0^{\tau} \frac{dt}{1+|B_t\!+\!x|}\bigg)
$$

for all stopping times τ of B , all $c > 1/2$ and all $x \ge 0$, where:

$$
\begin{aligned}
\text{(2.4.23)} \qquad W(x;c) \;&=\; \frac{1}{2c\!-\!1} + 2c\Big((1\!+\!x)\log(1\!+\!x)-x\Big) \quad \text{if} \;\; x \,\le\, 1/(2c\!-\!1) \\
&=\; 2c(1\!+\!x)\log\left(1\!+\!\frac{1}{2c\!-\!1}\right) - 1 \quad \text{if} \;\; x \,>\, 1/(2c\!-\!1) \;.
\end{aligned}
$$

This inequality is sharp, and for each $c \,>\, 1/2$ and $x \,\ge\, 0$ given and fixed, the equality in (2.4.23) is attained at the stopping time:

$$
\text{(2.4.24)} \qquad \tau_* = \inf\; \big\{\, t>0 \mid S_t - \alpha X_t \,\ge\, \beta \,\big\}
$$

which is pointwise the smallest possible with this property. By minimising over all $c > 1/2$ on the right-hand side in (2.4.22) we get a sharp inequality (the equality is attained at each stopping time τ_* from (2.4.24) whenever $c > 1/2$ and $x \ge 0$). In particular, this for $x = 0$ yields:

$$E\Big(\max_{0\le t\le\tau}|B_t|\Big) \le \frac{1}{2}\,E\Big(\int_0^\tau \frac{dt}{1+|B_t|}\Big) + \sqrt{2}\Big(E\int_0^\tau \frac{dt}{1+|B_t|}\Big)^{1/2} \tag{2.4.25}$$

for all stopping times τ of B. This inequality is sharp, and the equality in (2.4.25) is attained at each stopping time τ_* from (2.4.24). Note by Itô formula and the optional sampling theorem that:

$$E\Big(\int_0^\tau \frac{dt}{1+|B_t|}\Big) = 2E\Big(\big(1+|B_\tau|\big)\log\big(1+|B_\tau|\big)-|B_\tau|\Big) \tag{2.4.26}$$

for all stopping times τ of B satisfying $E(\tau^r)<\infty$ for some $r>1/2$. This shows that the inequality (2.4.25) in essence is of the $L\log L$-type. The advantage of (2.4.25) upon the classical Hardy-Littlewood $L\log L$-inequality is its sharpness for small stopping times as well (note that the equality in (2.4.25) is attained for $\tau\equiv 0$). For more information on this inequality and remaining details we refer to [P8].

Example 2.4.4 (***A sharp maximal inequality for geometric Brownian motion***) Consider the optimal stopping problem (2.2.4) where X is geometric Brownian motion and $c(x)\equiv c$. Recall that X is a non-negative diffusion having 0 as an entrance boundary point, and the infinitesimal operator of X in $]0,\infty[$ is given by the expression:

$$I\!L_X = \mu x\,\frac{\partial}{\partial x} + \frac{\sigma^2}{2}x^2\,\frac{\partial^2}{\partial x^2} \tag{2.4.27}$$

where $\mu\in I\!R$ and $\sigma>0$. The process X may be realised as:

$$X_t = x\exp\left(\sigma B_t + \Big(\mu-\frac{\sigma^2}{2}\Big)t\right) \tag{2.4.28}$$

with $x\ge 0$. The equation (2.3.21) takes the form:

$$g'(s) = \frac{\Delta\sigma^2 g^{\Delta+1}(s)}{2c\Big(s^\Delta - g^\Delta(s)\Big)} \tag{2.4.29}$$

where $\Delta = 1-2\mu/\sigma^2$. By using Picard's method of successive approximations it is possible to prove that for $\Delta>1$ the equation (2.4.29) admits the maximal solution $s\mapsto g_*(s)$ satisfying:

$$g_*(s)\sim s^{1-1/\Delta} \quad \text{(see } \textit{Figure}\text{ 2.2 below)} \tag{2.4.30}$$

for $s\to\infty$ (see [P11] for details). There seems to be no closed form for this solution. In the case $\Delta=1$ it is possible to find the general solution of (2.4.29) in a closed form, and this shows that the only non-negative solution is zero-function (see [P11]). By the result of Corollary 2.3.2 we may conclude that the payoff (2.2.4) is finite if and only if $\Delta>1$ (note that another argument was used in [P11] to obtain this equivalence), and in this case it is given by:

$$\begin{aligned} V_*(x,s) &= \frac{2c}{\Delta^2\sigma^2}\left(\left(\frac{x}{g_*(s)}\right)^\Delta - \log\left(\frac{x}{g_*(s)}\right)^\Delta - 1\right) + s \quad \text{if } g_*(s)<x\le s \\ &= s \quad \text{if } 0<x\le g_*(s)\,. \end{aligned} \tag{2.4.31}$$

The optimal stopping time is given by (2.3.57) with $s_* = 0$. By using explicit estimates from (2.4.30) on $s \to g_*(s)$ in (2.4.31), and then minimising over all $c > 0$, we obtain:

$$(2.4.32) \quad E\left(\max_{0 \le t \le \tau} \exp\left(\sigma B_t + \left(\mu - \frac{\sigma^2}{2} \right) t \right) \right) \le 1 - \frac{\sigma^2}{2\mu} + \frac{\sigma^2}{2\mu} \exp\left(- \frac{(\sigma^2 - 2\mu)^2}{2\sigma^2} E(\tau) - 1 \right)$$

for all stopping times τ of B . This inequality extends the well-known estimates of Doob in a sharp manner from deterministic times to stopping times. For more information and remaining details we refer to [P11]. Observe that the cost function $c(x) = c\,x$ in the optimal stopping problem (2.2.4) would imply that the maximal solution of (2.3.21) is linear. This shows that such a cost function suits better the maximum process and therefore is more natural. Explicit formulas for the payoff, and the maximal inequality obtained by minimising over $c > 0$, are also obtained easily in this case from the result of Corollary 2.3.2.

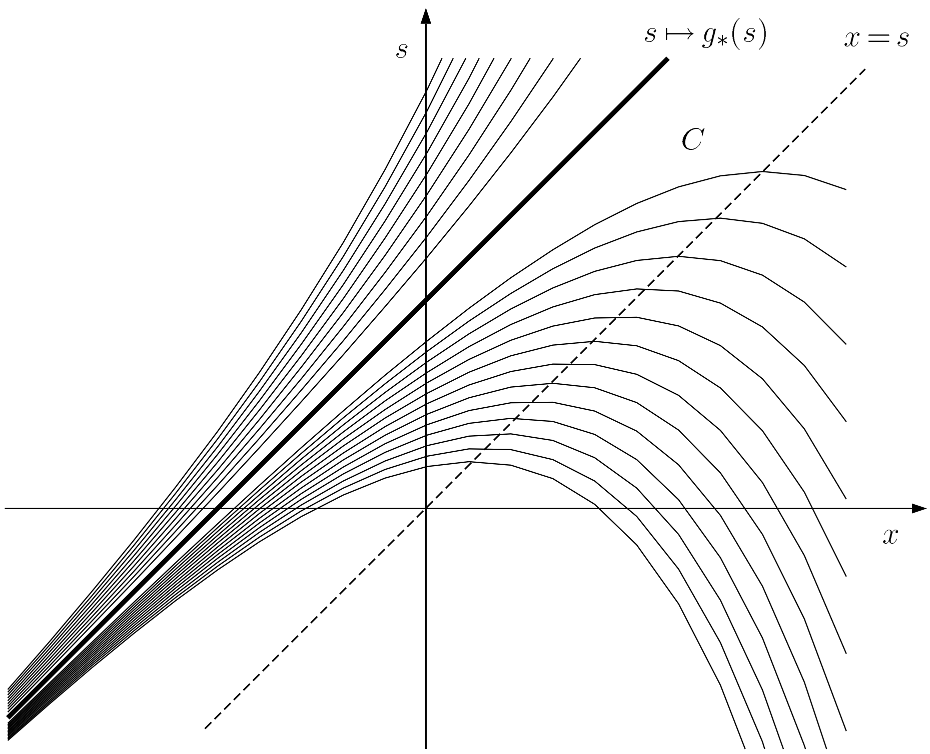


Figure 2.1. A computer drawing of solutions of the differential equation (2.3.21) in the case when $\mu \equiv 0$, $\sigma \equiv 1$ (thus $L(x) = x$) and $c \equiv 1/2$. The bold line $s \mapsto g_*(s)$ is the *maximal* solution which stays strictly below the diagonal in $I\!R^2$. (In this particular case $s \mapsto g_*(s)$ is a linear function.) By the maximality principle proved below, this solution is the optimal stopping boundary (the stopping time τ_* from (2.3.7) is optimal for the problem (2.2.4)).

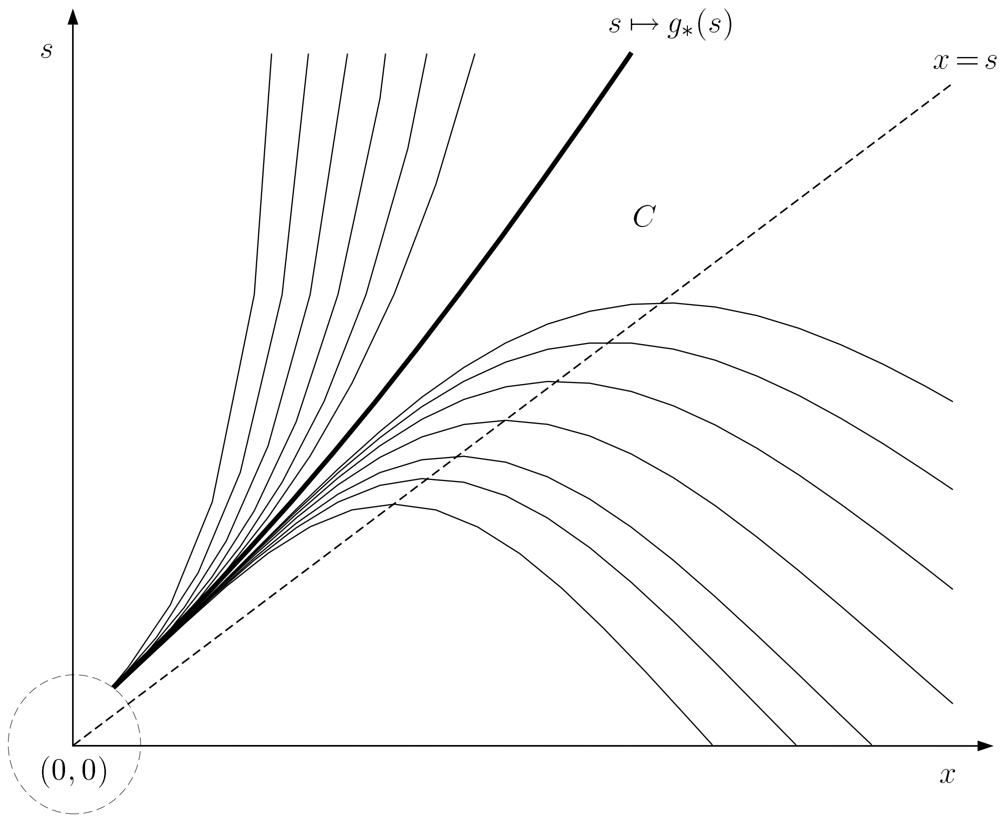


Figure 2.2. A computer drawing of solutions of the differential equation (2.3.21) in the case when X is a geometric Brownian motion from Example 2.4.4 below with $\mu = -1$, $\sigma^2 = 2$ (thus $\Delta = 2$) and $c = 50$. The bold line $s \mapsto g_*(s)$ is the *maximal* solution which stays strictly below the diagonal in $I\!R_+{}^2$. (In this particular case there is no closed formula for $s \mapsto g_*(s)$, but it is proved that $s \mapsto g_*(s)$ satisfies (2.4.30).)

3. Optimal prediction problems

Imagine the real-line movement of a Brownian particle started at 0 during the time interval $[0,1]$. Let S_1 denote the *maximal* positive height that the particle ever reaches during this time interval. As S_1 is a random quantity whose values depend on the entire Brownian path over the time interval, its *ultimate* value is at any given time $t \in [0,1)$ unknown. Following the Brownian particle from the initial time 0 onward, the question arises naturally as to determine a time when the movement should be terminated so that the position of the particle at that time is as 'close' as possible to the *ultimate maximum* S_1 . In this chapter we present the solution to this problem if 'closeness' is measured by a mean-square distance.

3.1 Formulation of the problem

To formulate the problem more precisely, let $B = (B_t)_{0 \le t \le 1}$ be a standard Brownian motion $(B_0 = 0, E(B_t) = 0, E(B_t^2) = t)$ defined on a probability space $(\Omega, \mathcal{F}, P)$, and let $I\!F^B = (\mathcal{F}_t^B)_{0 \le t \le 1}$ denote the natural filtration generated by B . Letting $\mathcal{M}$ denote the family of all stopping (Markov) times τ with respect to $I\!F^B$ satisfying $0 \le \tau \le 1$, the problem is to compute:

$$V_* = \inf_{\tau \in \mathcal{M}} E\left(B_\tau - \max_{0 \le t \le 1} B_t \right)^2 \tag{3.1.1}$$

and to find an optimal stopping time (the one at which the infimum in (3.1.1) is attained).

The solution of this problem is presented in Theorem 3.2.1 below. It turns out that the *maximum* process $S = (S_t)_{0 \le t \le 1}$ given by:

$$S_t = \sup_{0 \le s \le t} B_s \tag{3.1.2}$$

and the *CUSUM-type* reflected process $S - B = (S_t - B_t)_{0 \le t \le 1}$ plays a key role in the solution.

The optimal stopping problem (3.1.1) is of interest, for example, in financial mathematics and financial engineering where an optimal decision (i.e. optimal stopping time) should be based on a *prediction* of the future behaviour of the observable process (asset price, index, etc.). The argument also carries over to many other applied problems where such predictions play a role.

3.2 The result and proof

The main result of the chapter is contained in the next theorem. Below we let:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad \text{and} \quad \Phi(x) = \int_{-\infty}^{x} \varphi(y)\, dy \quad (x \in I\!R) \tag{3.2.1}$$

denote the density and distribution function of a standard normal variable.

Theorem 3.2.1

Consider the optimal stopping problem (3.1.1) where $(B_t)_{0 \le t \le 1}$ *is a standard Brownian motion. Then the value* V_* *is given by the formula:*

$$V_* = 2\Phi(z_*) - 1 = 0.73\ldots \tag{3.2.2}$$

where $z_* = 1.12\ldots$ *is the unique root of the equation:*

$$4\Phi(z_*) - 2z_*\varphi(z_*) - 3 = 0 \tag{3.2.3}$$

and the following stopping time is optimal (see Figures 3.2-3.5 below):

$$\tau_* = \inf \left\{ 0 \le t \le 1 \mid S_t - B_t \ge z_* \sqrt{1-t} \right\} \tag{3.2.4}$$

where S_t *is given by (3.1.2) above.*

Proof. Since $S_1 = \sup_{0 \le s \le 1} B_s$ is a square-integrable functional of the Brownian path on $[0,1]$, by the Itô-Clark representation theorem (see e.g. [57; p.191]) there exists a unique $I\!F^B$-adapted process $H = (H_t)_{0 \le t \le 1}$ satisfying $E(\int_0^1 H_t^2\, dt) < \infty$ such that:

$$S_1 = a + \int_0^1 H_t\, dB_t \tag{3.2.5}$$

where $a = E(S_1)$. Moreover, the following explicit formula is known to be valid:

$$H_t = 2\left(1 - \Phi\left(\frac{S_t - B_t}{\sqrt{1-t}}\right)\right) \tag{3.2.6}$$

for $0 \le t \le 1$ (see e.g. [58; p.93] and [44; p.365], or Section 3.3 below for a direct argument).

1. Associate with H the square-integrable martingale $M = (M_t)_{0 \le t \le 1}$ given by:

$$M_t = \int_0^t H_s\, dB_s\ . \tag{3.2.7}$$

By the martingale property of M and the optional sampling theorem, we obtain:

$$\begin{aligned} E(B_\tau - S_1)^2 &= E|B_\tau|^2 - 2E(B_\tau M_1) + E|S_1|^2 \\ &= E(\tau) - 2E(B_\tau M_\tau) + 1 = E\left(\int_0^\tau \big(1 - 2H_t\big)\, dt \right) + 1 \end{aligned} \tag{3.2.8}$$

for all $\tau \in \mathcal{M}$ (recall that $S_1 \sim |B_1|$). Inserting (3.2.6) into (3.2.8) we see that (3.1.1) reads:

$$V_* = \inf_{\tau \in \mathcal{M}} E\left(\int_0^\tau F\left(\frac{S_t - B_t}{\sqrt{1-t}}\right) dt \right) + 1 \tag{3.2.9}$$

where we denote $F(x) = 4\Phi(x) - 3$.

Since $S - B = (S_t - B_t)_{0 \le t \le 1}$ is a Markov process for which the natural filtration $I\!F^{S-B}$ coincides with the natural filtration $I\!F^B$, it follows from general theory of optimal stopping (see [65]) that in (3.2.9) we need only consider stopping times which are hitting times for $S - B$. Recalling moreover that $S - B \sim |B|$ by Lévy's distributional theorem (see e.g. [57; p.230]) and once more appealing to general theory, we see that (3.2.9) is equivalent to:

$$V_* = \inf_{\tau \in \mathcal{M}} E\left(\int_0^\tau F\left(\frac{|B_t|}{\sqrt{1-t}}\right) dt \right) + 1\ . \tag{3.2.10}$$

In our treatment of this problem, we first make use of a deterministic change of time.

2. Motivated by the form of (3.2.10), consider the process $Z=(Z_t)_{t\geq 0}$ given by:

$$Z_t = e^t \, B_{1-e^{-2t}} \; . \tag{3.2.11}$$

By Itô's formula we find that Z is a (strong) solution of the linear stochastic differential equation:

$$dZ_t = Z_t \, dt + \sqrt{2} \, d\beta_t \tag{3.2.12}$$

where the process $\beta = (\beta_t)_{0\leq t\leq 1}$ is given by:

$$\beta_t = \frac{1}{\sqrt{2}} \int_0^t e^s \, dB_{1-e^{-2s}} = \frac{1}{\sqrt{2}} \int_0^{1-e^{-2t}} \frac{1}{\sqrt{1-s}} \, dB_s \; . \tag{3.2.13}$$

As β is a continuous Gaussian martingale with mean zero and variance equal to t, it follows by Lévy's characterisation theorem (see e.g. [57; p.142]) that β is a standard Brownian motion. We thus may conclude that Z is a diffusion process with the infinitesimal operator given by:

$$\mathbb{L}_Z = z \frac{d}{dz} + \frac{d^2}{dz^2} \; . \tag{3.2.14}$$

Substituting $t = 1-e^{-2s}$ in (3.2.10) and using (3.2.11), we obtain:

$$V_* = 2 \inf_{\tau \in \mathcal{M}} E\left(\int_0^{\sigma_\tau} e^{-2s} F\big(|Z_s|\big) \, ds \right) + 1 \tag{3.2.15}$$

upon setting $\sigma_\tau = \log(1/\sqrt{1-\tau})$. It is clear from (3.2.11) that τ is a stopping time with respect to $\mathbb{F}^B$ if and only if σ_τ is a stopping time with respect to $\mathbb{F}^Z$. This shows that our initial problem (3.1.1) reduces to solve:

$$W_* = \inf_{\sigma} \, E\left(\int_0^{\sigma} e^{-2s} F\big(|Z_s|\big) \, ds \right) \tag{3.2.16}$$

where the infimum is taken over all $\mathbb{F}^Z$-stopping times σ with values in $[0,\infty]$. This problem belongs to the general theory of optimal stopping for time-homogeneous Markov processes (see [65]).

3. To calculate (3.2.16) define:

$$W_*(z) = \inf_{\sigma} \, E_z\left(\int_0^{\sigma} e^{-2s} F\big(|Z_s|\big) \, ds \right) \tag{3.2.17}$$

for $z \in \mathbb{R}$, where $Z_0 = z$ under P_z , and the infimum is taken as above. General theory combined with basic properties of the map $z \mapsto F(|z|)$ prompts that the stopping time:

$$\sigma_* = \inf \{ \, t > 0 \, : \, |Z_t| \geq z_* \, \} \tag{3.2.18}$$

should be optimal in (3.2.17), where $z_* > 0$ is a constant to be found.

To determine z_* and compute the value function $z \mapsto W_*(z)$ in (3.2.17), it is a matter of routine to formulate the following *free-boundary problem*:

$$(3.2.19)\qquad \big(I\!L_Z - 2\big)W(z) = -F(|z|) \quad \text{for} \quad z \in (-z_*, z_*)$$

$$(3.2.20)\qquad W(\pm z_*) = 0 \quad \text{(instantaneous stopping)}$$

$$(3.2.21)\qquad W'(\pm z_*) = 0 \quad \text{(smooth fit)}$$

where $I\!L_Z$ is given by (3.2.14) above. We shall extend the solution of (3.2.19)-(3.2.21) by setting its value equal to 0 for $z \notin (-z_*, z_*)$, and thus the map so obtained will be C^2 everywhere on $I\!R$ but at $-z_*$ and z_* where it is C^1.

Inserting $I\!L_Z$ from (3.2.14) into (3.2.19) leads to the following equation:

$$(3.2.22)\qquad W''(z) + zW'(z) - 2W(z) = -F(|z|)$$

for $z \in (-z_*, z_*)$. The form of the equation (3.2.12) and the value (3.2.16) indicates that $z \mapsto W_*(z)$ should be even; thus we shall additionally impose:

$$(3.2.23)\qquad W'(0) = 0$$

and consider (3.2.22) only for $z \in [0, z_*)$.

The general solution of the equation (3.2.22) for $z \geq 0$ is given by:

$$(3.2.24)\qquad W(z) = C_1(1+z^2) + C_2\big(z\varphi(z) + (1+z^2)\Phi(z)\big) + 2\Phi(z) - 3/2\ .$$

The three conditions $W(z_*) = W'(z_*) = W'(0) = 0$ determine constants C_1, C_2 and z_* uniquely; it is easily verified that $C_1 = \Phi(z_*)$, $C_2 = -1$, and z_* is the unique root of the equation (3.2.3). Inserting this back into (3.2.22), we obtain the following candidate for (3.2.17):

$$(3.2.25)\qquad W(z) = \Phi(z_*)(1+z^2) - z\varphi(z) + (1-z^2)\Phi(z) - 3/2$$

when $z \in [0, z_*]$, upon extending it to an even function on $I\!R$ as indicated above (see *Figure 3.1* below).

To verify that this solution $z \mapsto W(z)$ coincides with the value function (3.2.17), and that σ_* from (3.2.18) is an optimal stopping time, we shall note that $z \mapsto W(z)$ is C^2 everywhere but at $\pm z_*$ where it is C^1. Thus by the Itô-Tanaka formula we find:

$$(3.2.26)\qquad e^{-2t}W(Z_t) = W(Z_0) + \int_0^t e^{-2s}\Big(I\!L_Z W(Z_s) - 2\,W(Z_s)\Big)\,ds + \sqrt{2}\int_0^t e^{-2s}W'(Z_s)\,d\beta_s\ .$$

Hence by (3.2.22) and the fact that $I\!L_Z W(z) - 2\,W(z) = 0 > -F(|z|)$ for $z \notin [-z_*, z_*]$, upon extending W'' to $\pm z_*$ as we please and using that the Lebesgue measure of those $t > 0$ for which $Z_t = \pm z_*$ is zero, we get:

$$(3.2.27)\qquad e^{-2t}W(Z_t) \geq W(Z_0) - \int_0^t e^{-2s}F(|Z_s|)\,ds + M_t$$

where $M = (M_t)_{t\geq 0}$ is a continuous local martingale given by $M_t = \sqrt{2}\int_0^t e^{-2s}W'(Z_s)\,d\beta_s$.

Using further that $W(z) \leq 0$ for all z, a simple application of the optional sampling theorem in the stopped version of (3.2.27) under P_z shows that $W_*(z) \geq W(z)$ for all z. To prove equality one may note that the passage from (3.2.26) to (3.2.27) also yields:

$$0 = W(Z_0) - \int_0^{\sigma_*} e^{-2s} F(|Z_s|)\, ds + M_{\sigma_*} \tag{3.2.28}$$

upon using (3.2.19) and (3.2.20). Since clearly $E_z(\sigma_*) < \infty$ and thus $E_z(\sqrt{\sigma_*}) < \infty$ as well, and $z \mapsto W'(z)$ is bounded on $[-z_*, z_*]$, we can again apply the optional sampling theorem and conclude that $E_z(M_{\sigma_*}) = 0$. Taking the expectation under P_z on both sides in (3.2.28) enables one therefore to conclude $W_*(z) = W(z)$ for all z, and the proof of the claim is complete.

From (3.2.15)-(3.2.17) and (3.2.25) we find that $V_* = 2W_*(0) + 1 = 2(\Phi(z_*) - 1) + 1 = 2\Phi(z_*) - 1$. This establishes (3.2.2). Transforming σ_* from (3.2.18) back to the initial problem via the equivalence of (3.2.9), (3.2.10) and (3.2.15), we see that τ_* from (3.2.4) is optimal. The proof is complete. $\square$

Remarks:

1. Recalling that $S - B \sim |B|$ we see that τ_* is identically distributed as the stopping time $\widetilde{\tau} = \inf\{\, t > 0 : |B_t| = z_* \sqrt{1-t}\,\}$. This implies $E(\tau_*) = E(\widetilde{\tau}) = E|B_{\widetilde{\tau}}|^2 = (z_*)^2 E(1 - \widetilde{\tau}) = (z_*)^2 (1 - E(\tau_*))$, and hence we obtain:

$$E(\tau_*) = \frac{(z_*)^2}{1 + (z_*)^2} = 0.55\ldots \tag{3.2.29}$$

Moreover, using that $(B_t^4 - 6tB_t^2 + 3t^2)_{t \geq 0}$ is a martingale, similar arguments show that:

$$E(\tau_*)^2 = \frac{(z_*)^6 + 5(z_*)^4}{(1+(z_*)^2)(3+6(z_*)^2+(z_*)^4)} = 0.36\ldots \tag{3.2.30}$$

From (3.2.29) and (3.2.30) we find:

$$Var(\tau_*) = \frac{2(z_*)^4}{(1+(z_*)^2)^2(3+6(z_*)^2+(z_*)^4)} = 0.05\ldots \tag{3.2.31}$$

2. For the sake of comparison with (3.2.2) and (3.2.29) it is interesting to note that:

$$V_0 = \inf_{0 \leq t \leq 1} E\left(\Big(B_t - \max_{0 \leq s \leq 1} B_s \Big)^2 \right) = \frac{1}{\pi} + \frac{1}{2} = 0.81\ldots \tag{3.2.32}$$

with the infimum being attained at $t = 1/2$. For this, recall from (3.2.8) and (3.2.6) that:

$$E(B_t - S_1)^2 = E\left(\int_0^t F\left(\frac{S_s - B_s}{\sqrt{1-s}} \right) ds \right) + 1 \tag{3.2.33}$$

where $F(x) = 4\Phi(x) - 3$. Using further that $S - B \sim |B|$, elementary calculations show:

$$E(B_t - S_1)^2 = 4\left(\int_0^t E\left(\Phi\left(\frac{|B_s|}{\sqrt{1-s}} \right) \right) ds \right) - 3t + 1 \tag{3.2.34}$$

$$= 4 \int_0^t \left(1 - \frac{1}{\pi} \ \arctan \sqrt{\frac{1-s}{s}} \ \right) ds - 3t + 1$$

$$= -\frac{4}{\pi} \left(t \arctan \sqrt{\frac{1-t}{t}} + \frac{1}{2} \arctan \sqrt{\frac{t}{1-t}} - \frac{1}{2} \sqrt{t\,(1-t)} \ \right) + t + 1 \ .$$

Hence (3.2.32) is easily verified by standard means.

3. In view of the fact that σ_* from (3.2.18) with $z_* = 1.12\ldots$ from (3.2.3) is optimal in the problem (3.2.17), it is interesting to observe that the unique solution of the equation $F(\hat{z}) = 0$ is given by $\hat{z} = 0.67\ldots$ Noting moreover that the map $z \mapsto F(z)$ is increasing on $[0, \infty)$ and satisfies $F(0) = -1$, we see that $F(z) < 0$ for all $z \in [0, \hat{z})$ and $F(z) > 0$ for all $z > \hat{z}$. The size of the gap between $\hat{z}$ and z_* quantifies the tendency of the process $|Z|$ to return back to the 'favourable' region $[0, \hat{z})$ where clearly it is never optimal to stop.

4. The case of a general time interval $[0, T]$ easily reduces to the case of a unit time interval treated above by using the scaling property of Brownian motion implying:

$$\inf_{0 \le \tau \le T} E\left(\left(B_\tau - \max_{0 \le t \le T} B_t \right)^2 \right) = T \inf_{0 \le \tau \le 1} E\left(\left(B_\tau - \max_{0 \le t \le 1} B_t \right)^2 \right) \tag{3.2.35}$$

which further equals to $T(2\Phi(z_*) - 1)$ by (3.2.2). Moreover, the same argument shows that the optimal stopping time in (3.2.35) is given by:

$$\tau_* = \inf \left\{ \, 0 \le t \le T \ | \ S_t - B_t \ge z_* \sqrt{T - t} \ \right\} \tag{3.2.36}$$

where z_* is the same as in Theorem 3.2.1.

5. The maximum functional in the argument above can be replaced by other functionals. The integral functional is an example which turns out to have a trivial solution.

Setting $I_1 = \int_0^1 B_t \, dt$ we find by Itô's formula that the following analogue of (3.2.5) is valid:

$$I_1 = \int_0^1 (1 - t) \, dB_t \ . \tag{3.2.37}$$

Denoting $M_t = \int_0^t (1 - s) \, dB_s$ it follows as in (3.2.8) that:

$$E(B_\tau - I_1)^2 = E|B_\tau|^2 - 2E(B_\tau M_1) + E|I_1|^2 = E(\tau^2 - \tau) + 1/3 \tag{3.2.38}$$

for all $\tau \in \mathcal{M}$. Hence we see that (cf. (3.2.40) below):

$$\inf_{\tau \in \mathcal{M}} E(B_\tau - I_1)^2 = 1/12 = 0.08\ldots \tag{3.2.39}$$

and that the infimum is attained at $\tau_* \equiv 1/2$.

6. From the point of view of mathematical statistics, the "estimator" B_τ of S_1 is *biased*, since $E(B_\tau) = 0$ for all $0 \le \tau \le 1$ but $E(S_1) \ne 0$. It is thus desirable to consider the values:

$$\widetilde{V}_* = \inf_{a \in I\!R \, , \, \tau \in \mathcal{M}} E\Big(a + B_\tau - S_1 \Big)^2 \quad \text{and} \quad \widetilde{V}_0 = \inf_{a \in I\!R \, , \, 0 \le t \le 1} E\Big(a + B_t - S_1 \Big)^2 \tag{3.2.40}$$

and compare them with the values from (3.1.1) and (3.2.32). However, by using that $E(B_\tau) = 0$

we also find at once that $a_* = E(S_1)$ is optimal in (3.2.40) with $\widetilde{V}_* = V_* - 2/\pi = 0.09\ldots$ and $\widetilde{V}_0 = V_0 - 2/\pi = 0.18\ldots$

3.3 Appendix

In this section we present a *direct* derivation of the stochastic integral representation (3.2.5) and (3.2.6) (cf. [58; pp.89-93] and [44; pp.363-369]). For the sake of comparison we shall deal with a standard Brownian motion with drift given by:

$$B_t^\mu = B_t + \mu t \tag{3.3.1}$$

where μ is a real number. The maximum process S^μ associated with B^μ is given by:

$$S_t^\mu = \sup_{0 \le s \le t} B_s^\mu \; . \tag{3.3.2}$$

1. To derive the analogue of (3.2.5) and (3.2.6) in this case, we shall first note that *stationary independent increments* of B^μ imply:

$$\begin{aligned} E\big(S_1^\mu \,|\, \mathcal{F}_t^B\big) &= S_t^\mu + E\bigg(\Big(\sup_{t \le s \le 1} B_s^\mu - S_t^\mu\Big)^+ \,\Big|\, \mathcal{F}_t^B\bigg) \\ &= S_t^\mu + E\bigg(\Big(\sup_{t \le s \le 1} (B_s^\mu - B_t^\mu) - (S_t^\mu - B_t^\mu)\Big)^+ \,\Big|\, \mathcal{F}_t^B\bigg) \\ &= S_t^\mu + E\Big(S_{1-t}^\mu - (z-x)\Big)^+ \Big|_{z=S_t^\mu,\, x=B_t^\mu} \; . \end{aligned} \tag{3.3.3}$$

Using further the formula $E(X-c)^+ = \int_c^\infty P\{X>z\}\,dz$, we see that (3.3.3) reads as:

$$E\big(S_1^\mu \,|\, \mathcal{F}_t^B\big) = S_t^\mu + \int_{S_t^\mu - B_t^\mu}^\infty \big(1 - F_{1-t}^\mu(z)\big)\,dz := f(t, B_t^\mu, S_t^\mu) \tag{3.3.4}$$

where we use the following notation:

$$F_{1-t}^\mu(z) = P\big\{S_{1-t}^\mu \le z\big\} \tag{3.3.5}$$

and the map $f = f(t,x,s)$ is defined accordingly.

2. Applying Itô's formula to the right-hand side of (3.3.4), and using that the left-hand side defines a continuous martingale, we find upon setting $a_\mu = E(S_1^\mu)$ that:

$$\begin{aligned} E\big(S_1^\mu \,|\, \mathcal{F}_t^B\big) &= a_\mu + \int_0^t \frac{\partial f}{\partial x}(s, B_s^\mu, S_s^\mu)\,dB_s \\ &= a_\mu + \int_0^t \big(1 - F_{1-s}^\mu(S_s^\mu - B_s^\mu)\big)\,dB_s \end{aligned} \tag{3.3.6}$$

as a non-trivial continuous martingale cannot have paths of bounded variation. This reduces the initial problem to the problem of calculating (3.3.5).

3. The following explicit formula is well-known (see e.g. [44; p.368] or [66; pp.759-760]):

$$F^{\mu}_{1-t}(z) = \Phi\left(\frac{z-\mu(1-t)}{\sqrt{1-t}}\right) - e^{2\mu z}\,\Phi\left(\frac{-z-\mu(1-t)}{\sqrt{1-t}}\right) . \tag{3.3.7}$$

Inserting this into (3.3.6) we obtain the representation:

$$S^{\mu}_{1} = a_{\mu} + \int_0^1 H^{\mu}_t \, dB_t \tag{3.3.8}$$

where the process H^{μ} is explicitly given by:

$$H^{\mu}_t = 1 - \Phi\left(\frac{(S^{\mu}_t - B^{\mu}_t) - \mu(1-t)}{\sqrt{1-t}}\right) + e^{2\mu(S^{\mu}_t - B^{\mu}_t)}\,\Phi\left(\frac{-(S^{\mu}_t - B^{\mu}_t) - \mu(1-t)}{\sqrt{1-t}}\right) . \tag{3.3.9}$$

Setting $\mu = 0$ in this expression, we recover (3.2.5) and (3.2.6).

4. Note that the argument above extends to a large class of processes with stationary independent increments (including Lévy processes) by reducing the initial problem to calculating the analogue of (3.3.5). In particular, the following "prediction" result deserves a special note. It is derived in exactly the same way as (3.3.4) above.

Let $X = (X_t)_{0 \le t \le T}$ *be a process with stationary independent increments started at zero, and let us denote* $S_t = \max_{0 \le s \le t} X_s$ *for* $0 \le t \le T$ *. If* $E(S_T) < \infty$ *then the predictor* $E(S_T \,|\, \mathcal{F}^X_t)$ *of* S_T *based on the observations* $\{\, X_s \,|\, 0 \le s \le t \,\}$ *is given by the following formula:*

$$E\big(S_T \,|\, \mathcal{F}^X_t\big) = S_t + \int_{S_t - X_t}^{\infty} \big(1 - F_{T-t}(z)\big)\, dz \tag{3.3.10}$$

where $F_{T-t}(z) = P\{S_{T-t} \le z\}$.

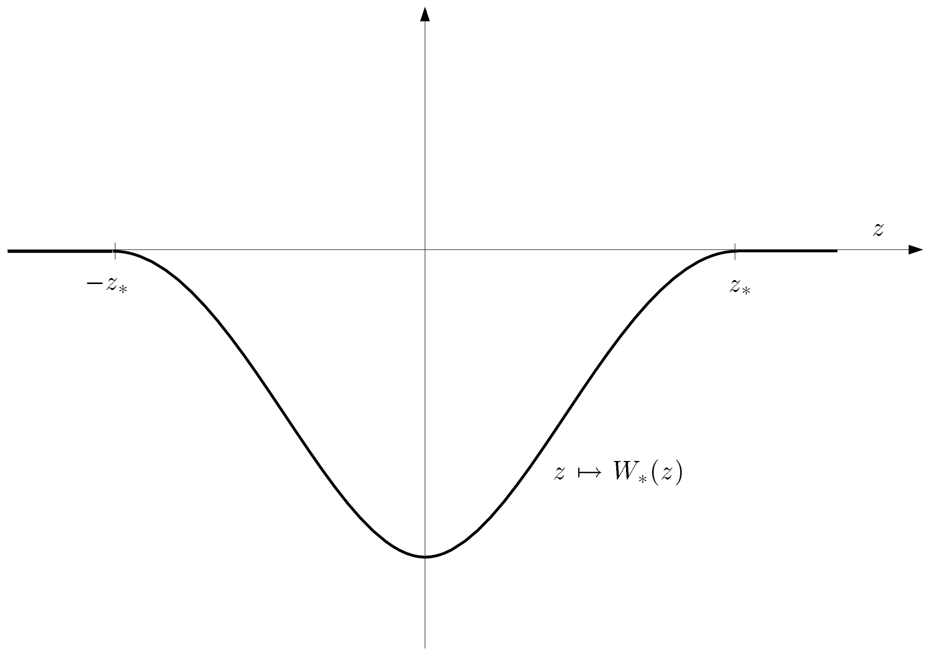


Figure 3.1. A computer drawing of the map (3.2.17). The smooth fit (3.2.21) holds at $-z_*$ and z_* .

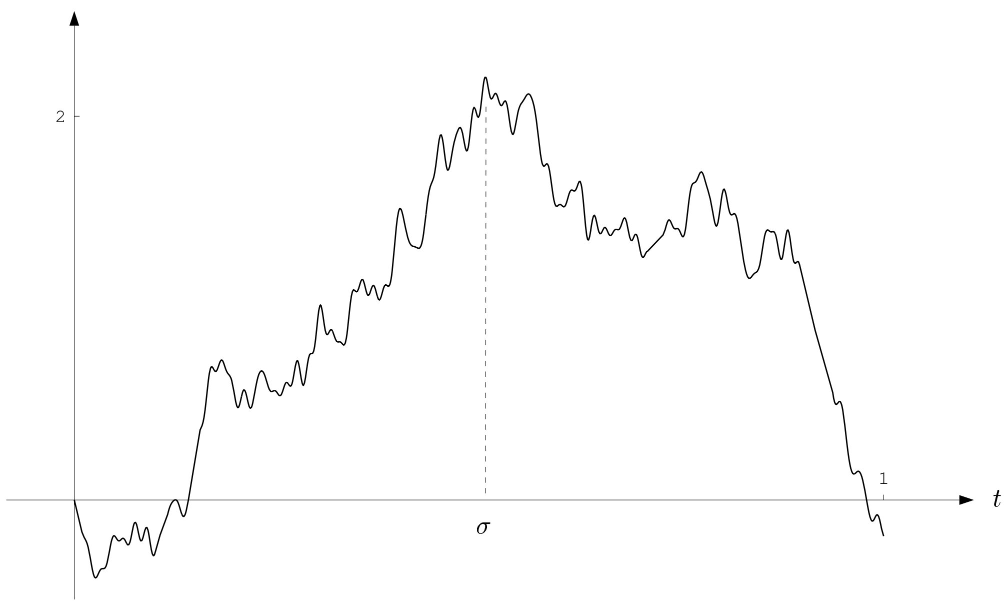


Figure 3.2. A computer simulation of a Brownian path $(B_t(\omega))_{0 \le t \le 1}$ with the maximum being attained at $\sigma = 0.51$.

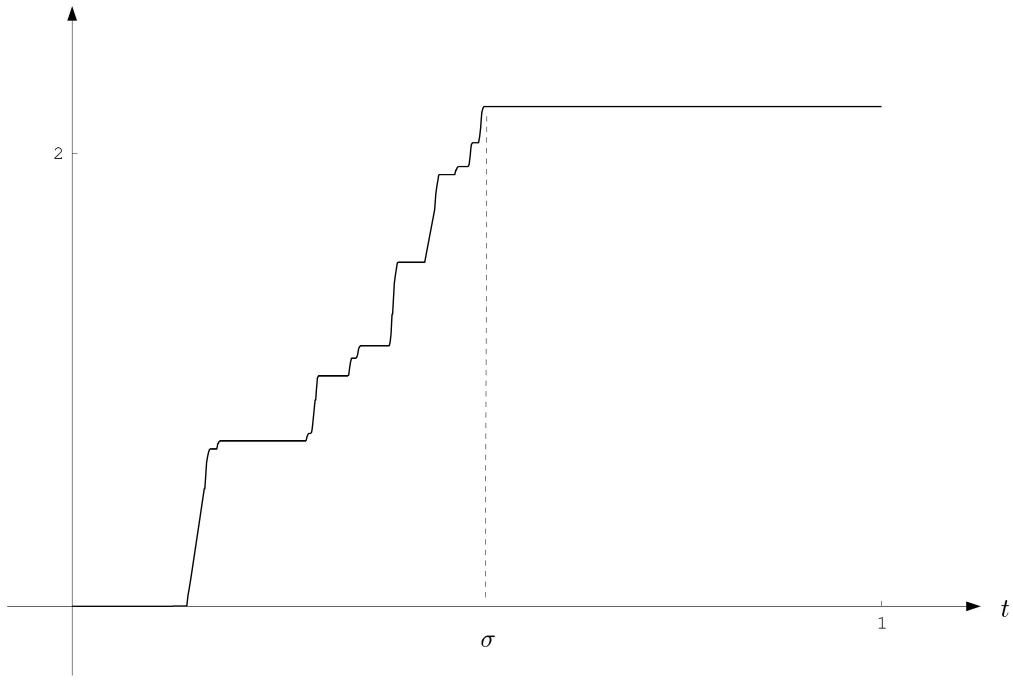


Figure 3.3. A computer drawing of the maximum process $(S_t(\omega))_{0 \le t \le 1}$ associated with the Brownian path from *Figure 3.2*.

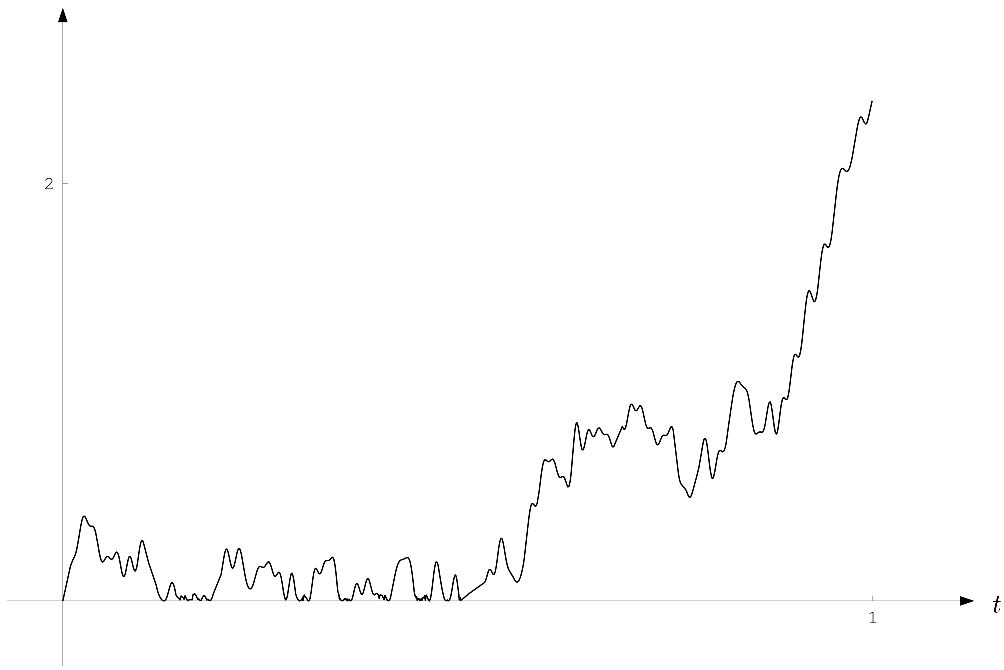


Figure 3.4. A computer drawing of the difference process $(S_t(\omega) - B_t(\omega))_{0 \le t \le 1}$ from *Figures 3.2-3.3*.

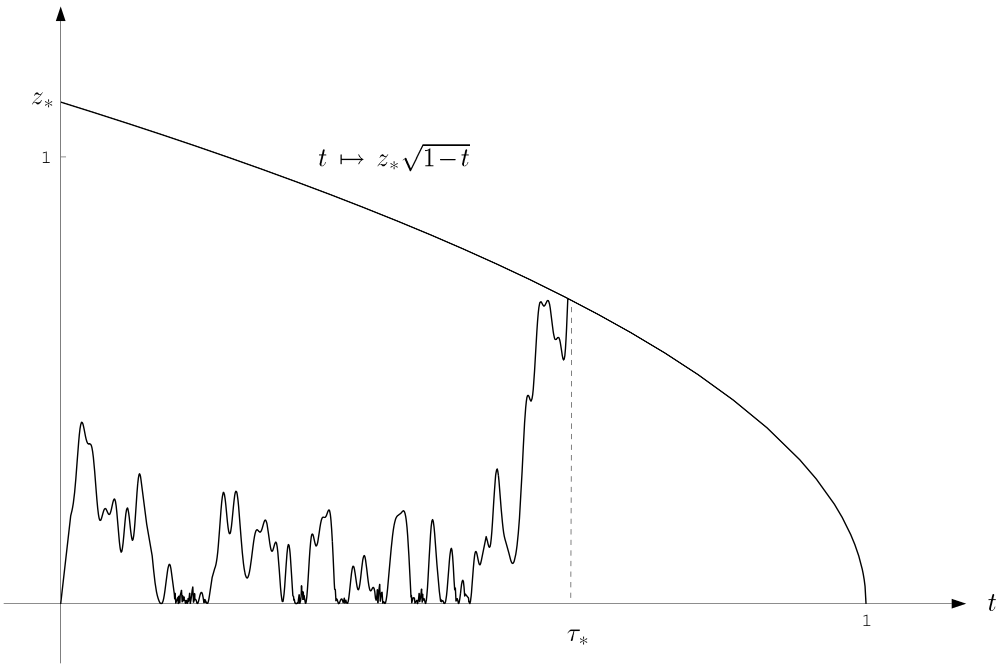


Figure 3.5. A computer drawing of of the optimal stopping strategy (3.2.4) for the Brownian path from *Figures 3.2-3.4*. It turns out that $\tau_* = 0.62$ in this case (cf. *Figure 3.2*).

4. Sequential testing problems

Suppose that at time $t=0$ we begin to observe a Poisson process $X = (X_t)_{t \geq 0}$ with intensity $\lambda > 0$ which is either λ_0 or λ_1 where $\lambda_0 < \lambda_1$. Assuming that the true value of λ is not known to us, our problem is then to decide as soon as possible and with a minimal error probability (both specified later) if the true value of λ is either λ_0 or λ_1 .

4.1 Description of the problem

Depending on the hypotheses about the unknown intensity λ , this problem admits two formulations. The *Bayesian* formulation relies upon the hypothesis that an a priori probability distribution of λ is given to us, and that λ takes either of the values λ_0 and λ_1 at time $t = 0$ according to this distribution. The *variational* formulation (sometimes also called a *fixed error probability* formulation) involves no probabilistic assumptions on the unknown intensity λ . The Wald sequential probability ratio test (SPRT) is known to be optimal in this context for a large class of observable processes (see [25, 38, 9]).

Despite the fact that the Bayesian approach to sequential analysis of problems on testing two statistical hypotheses has gained a considerable interest in the last fifty or so years (see e.g. [70, 71, 10, 47, 19, 65, 67]), it turns out that not many problems of that type have been solved explicitly (by obtaining a solution in closed form). In this respect the case of testing two simple hypotheses about the mean value of a Wiener process with drift is exceptional as the explicit solution to the problem has been obtained in both Bayesian and variational formulation. These solutions (including the proof of the optimality of the SPRT) were found by reducing the initial problem to a free-boundary problem (for a second order differential operator) which could be solved explicitly (see [64, 65]).

Our main aim in this chapter is to present the *explicit solution* of the Poisson intensity problem stated above in the context of a *Bayesian formulation* (Section 4.2), and then apply this result to deduce the optimality of the method (SPRT) in the context of a variational formulation (Section 4.3) with a precise description of the set of all admissible probabilities of a wrong decision ("errors of the first and second kind"). It will be clear from the sequel that the corresponding free-boundary problem becomes more delicate, since in the present case one needs to deal with a differential-difference operator, the appearance of which is a consequence of the discontinuous character of the observed (Poisson) process. The problem solved in Section 4.2 has been open for some time. (In the 1984 paper [38] the authors write that "in the case of Poisson processes, an explicit solution [of the Bayesian and free-boundary problem] is not known".)

4.2 Solution of the Bayesian problem

1. In the Bayesian formulation of the problem (see [65; Ch. 4]) it is assumed that at time $t = 0$ we begin observing a trajectory of the *point* process $X = (X_t)_{t \geq 0}$ with the compensator (see [49; Ch. 18]) $A = (A_t)_{t \geq 0}$, where $A_t = \lambda t$ and a random intensity $\lambda = \lambda(\omega)$ takes two values λ_1 and λ_0 with probabilities π and $1-\pi$. (We assume that $\lambda_1 > \lambda_0 > 0$ and $\pi \in [0,1]$.)

For a precise probability-statistical description of the Bayesian sequential testing problem it is convenient to assume that all our considerations take place on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P_\pi)$, where P_π has the following special structure:

(4.2.1) $$P_\pi = \pi P_1 + (1-\pi) P_0$$

for $\pi \in [0,1]$. We further assume that the $\mathcal{F}_0$-measurable random variable $\lambda = \lambda(\omega)$ takes two values λ_1 and λ_0 with probabilities $P_\pi(\lambda = \lambda_1) = \pi$ and $P_\pi(\lambda = \lambda_0) = 1 - \pi$. Concerning the observable point process $X = (X_t)_{t \geq 0}$, we assume that $P_\pi(X \in \cdot \mid \lambda = \lambda_i) = P_i(X \in \cdot\,)$, where $P_i(X \in \cdot\,)$ coincides with the distribution of a Poisson process with intensity λ_i for $i = 0, 1$.

Probabilities π and $1-\pi$ play a role of *a priori* probabilities of the statistical hypotheses:

(4.2.2) $$H_1 \,:\, \lambda = \lambda_1$$

(4.2.3) $$H_0 \,:\, \lambda = \lambda_0 \ .$$

2. Based upon the information which is continuously updated through the observation of the point process X , our problem is to test sequentially the hypotheses H_1 and H_0 . For this it is assumed that we have at disposal a class of sequential *decision rules* (τ, d) consisting of *stopping times* $\tau = \tau(\omega)$ with respect to $(\mathcal{F}_t^X)_{t \geq 0}$ where $\mathcal{F}_t^X = \sigma\{X_s \mid s \leq t\}$, and $\mathcal{F}_\tau^X$-measurable functions $d = d(\omega)$ which take values 0 and 1 . Stopping the observation of X at time τ , the *terminal decision function* d indicates that either the hypothesis H_1 or the hypothesis H_0 should be accepted; if $d = 1$ we accept H_1 , and if $d = 0$ we accept that H_0 is true.

3. Each decision rule (τ, d) implies losses of two kinds: the loss due to a cost of the observation, and the loss due to a wrong terminal decision. The average loss of the first kind may be naturally identified with $cE_\pi(\tau)$, and the average loss of the second kind can be expressed as $a\, P_\pi(d=0, \lambda=\lambda_1) + b\, P_\pi(d=1, \lambda=\lambda_0)$, where $c, a, b > 0$ are some constants. It will be clear from (4.2.8) below that there is no restriction to assume that $c = 1$, as the case of general $c > 0$ follows by replacing a and b with a/c and b/c respectively. Thus, the *total average loss* of the decision rule (τ, d) is given by:

(4.2.4) $$L_\pi(\tau, d) = E_\pi\Big(\, \tau + a\, 1_{(d=0\,,\lambda=\lambda_1)} + b\, 1_{(d=1\,,\lambda=\lambda_0)} \Big) \ .$$

Our problem is then to compute:

(4.2.5) $$V(\pi) = \inf_{(\tau,d)} L_\pi(\tau, d)$$

and to find the optimal decision rule (τ_*, d_*) , called *the π-Bayes decision rule*, at which the infimum in (4.2.5) is attained.

Observe that for any decision rule (τ, d) we have:

(4.2.6) $$a\, P_\pi(d=0, \lambda=\lambda_1) \;+\; b\, P_\pi(d=1, \lambda=\lambda_0) \;=\; a\, \pi\, \alpha(d) + b\, (1-\pi)\, \beta(d)$$

where $\alpha(d) = P_1(d=0)$ is called *the probability of an error of the first kind*, and $\beta(d) = P_0(d=1)$ is called *the probability of an error of the second kind.*

4. The problem (4.2.5) can be reduced to an optimal stopping problem for the *a posteriori probability process* defined by:

(4.2.7) $$\pi_t = P_\pi\big(\lambda = \lambda_1 \mid \mathcal{F}_t^X\big)$$

with $\pi_0 = \pi$ under P_π . Standard arguments (see [65; pp.166-167]) show that:

$$V(\pi) \;=\; \inf_\tau \; E_\pi\big(\, \tau + g_{a,b}(\pi_\tau)\,\big) \tag{4.2.8}$$

where $g_{a,b}(\pi) = a\pi \wedge b(1-\pi)$; the optimal stopping time τ_* in (4.2.8) is also optimal in (4.2.5), and the optimal decision function d_* is obtained by setting:

$$\begin{aligned} d_* \;&=\; 1 \quad \text{if} \quad \pi_{\tau_*} \geq b/(a+b) \\ &=\; 0 \quad \text{if} \quad \pi_{\tau_*} < b/(a+b)\ . \end{aligned} \tag{4.2.9}$$

Our main task in the sequel is therefore reduced to solving the optimal stopping problem (4.2.8).

5. Another natural process, which is in a one-to-one correspondence with the process $(\pi_t)_{t\geq 0}$, is *the likelihood ratio process*; it is defined as the Radon-Nikodym density:

$$\varphi_t \;=\; \frac{d\,(P_1\,|\,\mathcal{F}_t^X)}{d\,(P_0\,|\,\mathcal{F}_t^X)} \tag{4.2.10}$$

where $P_i\,|\,\mathcal{F}_t^X$ denotes the restriction of P_i to $\mathcal{F}_t^X$ for $i = 0,1$. Since:

$$\pi_t \;=\; \pi\,\frac{d\,(P_1\,|\,\mathcal{F}_t^X)}{d\,(P_\pi\,|\,\mathcal{F}_t^X)} \tag{4.2.11}$$

where $P_\pi \mid \mathcal{F}_t^X \;=\; \pi\, P_1 \mid \mathcal{F}_t^X + (1-\pi)\, P_0 \mid \mathcal{F}_t^X$, it follows that:

$$\pi_t \;=\; \left(\frac{\pi}{1-\pi}\;\varphi_t\right)\bigg/\left(1+\frac{\pi}{1-\pi}\;\varphi_t\right) \tag{4.2.12}$$

as well as that:

$$\varphi_t \;=\; \frac{1-\pi}{\pi}\;\frac{\pi_t}{1-\pi_t}\ . \tag{4.2.13}$$

Moreover, the explicit expression is known to be valid (see e.g. [25] or [49; Theorem 19.7]):

$$\varphi_t \;=\; \exp\left(X_t \log\left(\frac{\lambda_1}{\lambda_0}\right) - (\lambda_1 - \lambda_0)\, t\right)\ . \tag{4.2.14}$$

This representation may now be used to reveal the *Markovian* structure in the problem. Since the process $(X_t)_{t\geq 0}$ is a time-homogeneous Markov process having stationary independent increments (Lévy process) under both P_0 and P_1 , from the representation (4.2.14), and due to the one-to-one correspondence (4.2.12), we see that $(\varphi_t)_{t\geq 0}$ and $(\pi_t)_{t\geq 0}$ are time-homogeneous Markov processes under both P_0 and P_1 with respect to natural filtrations which clearly coincide with $(\mathcal{F}_t^X)_{t\geq 0}$. Using then further that $E_\pi(Y \mid \mathcal{F}_t^X) \;=\; E_1(Y \mid \mathcal{F}_t^X)\,\pi_t + E_0(Y \mid \mathcal{F}_t^X)\,(1-\pi_t)$ for any (bounded) measurable Y , it follows that $(\pi_t)_{t\geq 0}$, and thus $(\varphi_t)_{t\geq 0}$ as well, is a time-homogeneous Markov process under each P_π for $\pi \in [0,1]$. (Observe, however, that although the same argument shows that $(X_t)_{t\geq 0}$ is a Markov process under each P_π for $\pi \in \big\langle 0,1 \big\rangle$, it is not a time-homogeneous Markov process unless π equals 0 or 1 .) Note also directly from (4.2.7) that $(\pi_t)_{t\geq 0}$ is a martingale under each P_π for $\pi \in [0,1]$. Thus, the optimal stopping problem (4.2.8) falls into the class of optimal stopping problems for Markov processes, and we therefore proceed by finding the infinitesimal operator of $(\pi_t)_{t\ge 0}$. A slight modification of the arguments above shows that all these processes possess a strong Markov property actually.

6. By Itô's formula (see e.g. [50; Ch. 2, §3] or [41; Ch. I, §4]) one can verify that processes $(\varphi_t)_{t\ge 0}$ and $(\pi_t)_{t\ge 0}$ solve the following stochastic equations respectively:

$$d\varphi_t = \left(\frac{\lambda_1}{\lambda_0} - 1\right)\varphi_{t-}\, d\big(X_t - \lambda_0 t\big) \tag{4.2.15}$$

$$d\pi_t = \frac{(\lambda_1 - \lambda_0)\,\pi_{t-}(1-\pi_{t-})}{\lambda_1\pi_{t-} + \lambda_0\,(1-\pi_{t-})}\left(dX_t - \Big(\lambda_1\pi_{t-} + \lambda_0\,(1-\pi_{t-})\Big)\,dt\right) \tag{4.2.16}$$

(cf. formula (19.86) in [49]). The equation (4.2.16) may now be used to determine the infinitesimal operator of the Markov process $(\pi_t, \mathcal{F}_t^X, P_\pi)_{t\ge 0}$ for $\pi \in [0,1]$. For this, let $f = f(\pi)$ from $C^1[0,1]$ be given. Then by Itô's formula we find:

$$\begin{aligned}
(4.2.17)\quad f(\pi_t) &= f(\pi_0) + \int_0^t f'(\pi_{s-})\, d\pi_s + \sum_{0<s\le t}\Big(f(\pi_s) - f(\pi_{s-}) - f'(\pi_{s-})\,\Delta\pi_s\Big)\\
&= f(\pi_0) + \int_0^t f'(\pi_{s-})\Big(-(\lambda_1-\lambda_0)\,\pi_{s-}(1-\pi_{s-})\Big)\,ds + \sum_{0<s\le t}\Big(f(\pi_s) - f(\pi_{s-})\Big)\\
&= f(\pi_0) + \int_0^t f'(\pi_{s-})\Big(-(\lambda_1-\lambda_0)\pi_{s-}(1-\pi_{s-})\Big)\,ds + \int_0^t\int_0^1\Big(f(\pi_{s-}+y) - f(\pi_{s-})\Big)\,\mu^\pi(ds,dy)\\
&= f(\pi_0) + \int_0^t f'(\pi_{s-})\Big(-(\lambda_1-\lambda_0)\pi_{s-}(1-\pi_{s-})\Big)\,ds + \int_0^t\int_0^1\Big(f(\pi_{s-}+y) - f(\pi_{s-})\Big)\,\nu^\pi(ds,dy)\\
&\quad + \int_0^t\int_0^1\Big(f(\pi_{s-}+y) - f(\pi_{s-})\Big)\Big(\mu^\pi(ds,dy) - \nu^\pi(ds,dy)\Big)\\
&= f(\pi_0) + \int_0^t (\mathbb{L}f)(\pi_{s-})\,ds + M_t
\end{aligned}$$

where μ^π is the random measure of jumps of the process $(\pi_t)_{t\ge 0}$ and ν^π is a compensator of μ^π (see e.g. [50; Ch. 3] or [41; Ch. II]), the operator $\mathbb{L}$ is given as in (4.2.19) below, and $M = (M_t)_{t\ge 0}$ defined as:

$$M_t = \int_0^t\int_0^1\Big(f(\pi_{s-}+y) - f(\pi_{s-})\Big)\Big(\mu^\pi(ds,dy) - \nu^\pi(ds,dy)\Big) \tag{4.2.18}$$

is a local martingale with respect to $(\mathcal{F}_t^X)_{t\ge 0}$ and P_π for every $\pi \in [0,1]$. It follows from (4.2.17) that the infinitesimal operator of $(\pi_t)_{t\ge 0}$ acts on $f \in C^1[0,1]$ like:

$$(\mathbb{L}f)(\pi) = -(\lambda_1-\lambda_0)\pi(1-\pi)f'(\pi) + \Big(\lambda_1\pi + \lambda_0(1-\pi)\Big)\left(f\left(\frac{\lambda_1\pi}{\lambda_1\pi + \lambda_0\,(1-\pi)}\right) - f(\pi)\right). \tag{4.2.19}$$

7. Looking back at (4.2.5) and using explicit expressions (4.2.4) and (4.2.6) with (4.2.1), it is easily verified (cf. [47; p.105]) that the payoff $\pi \mapsto V(\pi)$ is a concave function on $[0,1]$, and thus it is continuous on $\langle 0,1\rangle$. Evidently, this function is pointwise dominated by $\pi \mapsto g_{a,b}(\pi)$. From these facts and from the general theory of optimal stopping for Markov processes (see e.g. [65])

we may guess that the payoff $\pi \mapsto V(\pi)$ from (4.2.8) should solve the following *free-boundary problem* (for a differential-difference equation defined by the infinitesimal operator):

$$(\mathbb{L}V)(\pi) = -1 \; , \quad A_* < \pi < B_* \tag{4.2.20}$$

$$V(\pi) = a\,\pi \wedge b\,(1-\pi) \; , \quad \pi \notin \langle A_*, B_* \rangle \tag{4.2.21}$$

$$V(A_*+) = V(A_*) \; , \; V(B_*-) = V(B_*) \quad \text{(continuous fit)} \tag{4.2.22}$$

$$V'(A_*) = a \quad \text{(smooth fit)} \tag{4.2.23}$$

for some $0 < A_* < b/(a+b) < B_* < 1$ to be found. Observe that (4.2.21) contains two conditions relevant for the system: (i) $V(A_*) = aA_*$ and (ii) $V(\pi) = b(1-\pi)$ for $\pi \in [B_*, S(B_*)]$ with $S = S(\pi)$ from (4.2.24) below. These conditions are in accordance with the fact that if the process $(\pi_t)_{t \geq 0}$ starts or ends up at some π outside $\langle A_*, B_* \rangle$, we must stop it instantly.

Note from (4.2.16) that the process $(\pi_t)_{t \geq 0}$ moves continuously towards 0 and only jumps towards 1 at times of jumps of the point process X. This provides some intuitive support for the principle of smooth fit to hold at A_*. However, without a concavity argument it is not a priori clear why the condition $V(B_*-) = V(B_*)$ should hold at B_*. As *Figure* 4.1 below shows, this is a rare property shared only by exceptional pairs (A, B), and one could think that once A_* is fixed through the "smooth fit", the unknown B_* will be determined uniquely through the "continuous fit". While this train of thoughts sounds perfectly logical, we shall see quite opposite below that the equation (4.2.19) dictates our travel to solution from B_* to A_*.

Our next aim is to show that the three conditions in (4.2.22) and (4.2.23) are sufficient to determine a unique solution of the free-boundary problem which in turn leads to the solution of the optimal stopping problem (4.2.8).

8. *Solution of the free-boundary problem (4.2.20)-(4.2.23).* Consider the equation (4.2.20) on $\langle 0, B]$ with some $B > b/(a+b)$ given and fixed. Introduce the "step" function:

$$S(\pi) = \frac{\lambda_1\,\pi}{\lambda_1\,\pi + \lambda_0\,(1-\pi)} \tag{4.2.24}$$

for $\pi \leq B$. Observe that $\pi \mapsto S(\pi)$ is increasing, and find points $\ldots < B_2 < B_1 < B_0 := B$ such that $S(B_n) = B_{n-1}$ for $n \geq 1$. It is easily verified that:

$$B_n = \frac{(\lambda_0)^n B}{(\lambda_0)^n B + (\lambda_1)^n (1-B)} \quad (n = 0, 1, \ldots\,) \,. \tag{4.2.25}$$

Denote $I_n = \langle B_n, B_{n-1}]$ for $n \geq 1$, and introduce the "distance" function:

$$d(\pi, B) = 1 + \left[\log\left(\frac{B}{1-B}\,\frac{1-\pi}{\pi} \right) \Big/ \log\left(\frac{\lambda_1}{\lambda_0} \right) \right] \tag{4.2.26}$$

for $\pi \leq B$, where $[x]$ denotes the integer part of x. Observe that d is defined to satisfy:

$$\pi \in I_n \iff d(\pi, B) = n \tag{4.2.27}$$

for all $0 < \pi \leq B$.

Consider the equation (4.2.20) on I_1 upon setting $V(\pi) = b(1-\pi)$ for $\pi \in \langle B, S(B)]$; this is then a first-order linear differential equation which can be solved explicitly, and imposing a continuity condition at B which is in agreement with (4.2.22), we obtain a *unique* solution $\pi \mapsto V(\pi;B)$ on I_1 ; move then further and consider the equation (4.2.20) on I_2 upon using the solution found on I_1 ; this is then a first-order linear differential equation which can be solved explicitly, and imposing a continuity condition over $I_2 \cup I_1$ at B_1 , we obtain a *unique* solution $\pi \mapsto V(\pi;B)$ on I_2 ; continuing this process by induction, we find the following formula:

$$\text{(4.2.28)} \qquad V(\pi;B) = \frac{(1-\pi)^{\gamma_1}}{\pi^{\gamma_0}} \sum_{k=0}^{n-1} \left(C_{n-k}\, \frac{\beta^k}{k!}\, \log^k \left(\left(\frac{\lambda_1}{\lambda_0} \right)^{k-1} \frac{\pi}{1-\pi} \right) \right) - \left(n\, \frac{\lambda_1-\lambda_0}{\lambda_0\lambda_1} + b \right) \pi + \left(\frac{n}{\lambda_0} + b \right)$$

for $\pi \in I_n$, where $C_1, \ldots, C_n$ are constants satisfying the following recurrent relation:

$$\text{(4.2.29)} \qquad C_{p+1} = \sum_{k=0}^{p-1} \left(C_{p-k} \left(f_k^{(p)} - f_{k+1}^{(p)} \right) \right) + \frac{(B_p)^{\gamma_0}}{(1-B_p)^{\gamma_1}} \left(\frac{\lambda_1-\lambda_0}{\lambda_0\lambda_1}\, B_p - \frac{1}{\lambda_0} \right)$$

for $p = 0, 1, \ldots, n-1$, with:

$$\text{(4.2.30)} \qquad f_k^{(p)} = \frac{\beta^k}{k!}\, \log^k \left(\left(\frac{\lambda_1}{\lambda_0} \right)^{k-p-1} \frac{B}{1-B} \right)$$

and where we set:

$$\text{(4.2.31)} \qquad \gamma_0 = \frac{\lambda_0}{\lambda_1-\lambda_0} \; ; \; \gamma_1 = \frac{\lambda_1}{\lambda_1-\lambda_0} \; ; \; \beta = \frac{1}{(\lambda_1-\lambda_0)}\, \frac{(\lambda_0)^{\gamma_1}}{(\lambda_1)^{\gamma_0}} \, .$$

Making use of the distance function (4.2.26), we may now write down the unique solution of (4.2.20) on $\langle 0, B]$ satisfying (4.2.21) on $[B, S(B)]$ and the second part of (4.2.22) at B :

$$\text{(4.2.32)} \qquad V(\pi;B) = \frac{(1-\pi)^{\gamma_1}}{\pi^{\gamma_0}} \sum_{k=0}^{d(\pi,B)-1} \left(C_{d(\pi,B)-k}\, \frac{\beta^k}{k!}\, \log^k \left(\left(\frac{\lambda_1}{\lambda_0} \right)^{k-1} \frac{\pi}{1-\pi} \right) \right) - \left(d(\pi,B)\, \frac{\lambda_1-\lambda_0}{\lambda_0\lambda_1} + b \right) \pi + \left(\frac{d(\pi,B)}{\lambda_0} + b \right)$$

for $0 < \pi \le B$. It is clear from our construction above that $\pi \mapsto V(\pi;B)$ is C^1 on $\langle 0, B \rangle$ and C^0 at B .

Observe that when computing the first derivative of $\pi \mapsto V(\pi;B)$, we can treat $d(\pi,B)$ in (4.2.32) as not depending on π . This then gives the following explicit expression:

$$\text{(4.2.33)} \qquad V'(\pi;B) = \frac{(1-\pi)^{\gamma_1-1}}{\pi^{\gamma_0+1}} \sum_{k=0}^{d(\pi,B)-1} \left(C_{d(\pi,B)-k}\, \frac{\beta^k}{k!}\, \log^k \left(\left(\frac{\lambda_1}{\lambda_0} \right)^{k-1} \frac{\pi}{1-\pi} \right) \cdot \right.$$

$$\cdot \left(k \Big/ \log \left(\left(\frac{\lambda_1}{\lambda_0} \right)^{k-1} \frac{\pi}{1-\pi} \right) - (\pi + \gamma_0) \right) \Bigg) - \left(d(\pi, B) \, \frac{\lambda_1 - \lambda_0}{\lambda_0 \lambda_1} + b \right)$$

for $0 < \pi \leq B$.

Setting $C = b/(a+b)$ elementary calculations show that $\pi \mapsto V(\pi; B)$ is concave on $\langle 0, B \rangle$, as well as that $V(\pi; B) \to -\infty$ as $\pi \downarrow 0$, for all $B \in [C, 1]$. Moreover, it is easily seen from (4.2.28) (with $n = 1$) that $V(\pi; 1) < 0$ for all $0 < \pi < 1$. Thus, if for some $\hat{B} > C$, close to C , it happens that $\pi \mapsto V(\pi; \hat{B})$ crosses $\pi \mapsto a\pi$ when π moves to the left from $\hat{B}$, then a uniqueness argument presented in Remark 4.2.2 below (for different B's the curves $\pi \mapsto V(\pi; B)$ do not intersect) shows that there exists $B_* \in \langle C, 1 \rangle$, obtained by moving B from $\hat{B}$ to 1 or vice versa, such that for some $A_* \in \langle 0, C \rangle$ we have $V(A_*; B_*) = a\, A_*$ and $V'(A_*; B_*) = a$ (see *Figure* 4.2 below). Observe that the first identity captures part (i) of (4.2.22), while the second settles (4.2.23).

These considerations show that the system (4.2.20)-(4.2.23) has a unique (non-trivial) solution consisting of A_* , B_* and $\pi \mapsto V(\pi; B_*)$, if and only if:

(4.2.34) $$\lim_{B \downarrow C} V'(B-; B) < a \ .$$

Geometrically this is the case when for $B > C$, close to C , the solution $\pi \mapsto V(\pi; B)$ intersects $\pi \mapsto a\pi$ at some $\pi < B$. It is now easily verified by using (4.2.28) (with $n = 1$) that (4.2.34) holds if and only if the following condition is satisfied:

(4.2.35) $$\lambda_1 - \lambda_0 > \frac{1}{a} + \frac{1}{b} \ .$$

In this process one should observe that B_1 from (4.2.25) tends to a number strictly less than C when $B \downarrow C$, so that all calculations are actually performed on I_1 .

Thus, the condition (4.2.35) is necessary and sufficient for the existence of a unique non-trivial solution of the system (4.2.20)-(4.2.23); in this case the optimal A_* and B_* are uniquely determined as the solution of the system of transcendental equations $V(A_*; B_*) = a\, A_*$ and $V'(A_*; B_*) = a$, where $\pi \mapsto V(\pi; B)$ and $\pi \mapsto V'(\pi; B)$ are given by (4.2.32) and (4.2.33) respectively; once A_* and B_* are fixed, the solution $\pi \mapsto V(\pi; B_*)$ is given for $\pi \in [A_*, B_*]$ by means of (4.2.32).

9. *Solution of the optimal stopping problem (4.2.8).* We shall now show that the solution of the free-boundary problem (4.2.20)-(4.2.23) found above coincides with the solution of the optimal stopping problem (4.2.8). This in turn leads to the solution of the Bayesian problem (4.2.5).

Theorem 4.2.1

(I): Suppose that the condition (4.2.35) holds. Then the π-Bayes decision rule (τ_*, d_*) in the problem (4.2.5) of testing two simple hypotheses H_1 and H_0 is explicitly given by:

(4.2.36) $$\tau_* = \inf \left\{ t \geq 0 \mid \pi_t \notin \langle A_*, B_* \rangle \right\}$$

(4.2.37) $$d_* = 1 \;\; \textit{(accept } H_1 \textit{)}, \;\; \textit{if} \;\; \pi_{\tau_*} \geq B_*$$

$$= 0 \quad (\,accept\ \ H_0\,)\,, \quad if \quad \pi_{\tau_*} = A_*$$

(see Remark 4.2.3) where the constants A_* *and* B_* *satisfying* $0 < A_* < b/(a{+}b) < B_* < 1$ *are uniquely determined as solutions of the system of transcendental equations:*

$$V(A_*; B_*) \;=\; a A_* \tag{4.2.38}$$

$$V'(A_*; B_*) \;=\; a \tag{4.2.39}$$

with $\pi \mapsto V(\pi; B)$ *and* $\pi \mapsto V'(\pi; B)$ *in (4.2.32) and (4.2.33) respectively.*

(II): In the case when the condition (4.2.35) fails to hold, the π*-Bayes decision rule is trivial: Accept* H_1 *if* $\pi > b/(a{+}b)$ *, and accept* H_0 *if* $\pi < b/(a{+}b)$ *; either decision is equally good if* $\pi = b/(a{+}b)$ *.*

Proof. (I): 1. We showed above that the free-boundary problem (4.2.20)-(4.2.23) is solvable if and only if (4.2.35) holds, and in this case the solution $\pi \mapsto V_*(\pi)$ is given explicitly by $\pi \mapsto V(\pi; B_*)$ in (4.2.32) for $A_* \le \pi \le B_*$, where A_* and B_* are a unique solution of (4.2.38)-(4.2.39).

In accordance with the interpretation of the free-boundary problem, we extend $\pi \mapsto V_*(\pi)$ to the whole of $[0,1]$ by setting $V_*(\pi) = a\,\pi$ for $0 \le \pi < A_*$ and $V_*(\pi) = b\,(1{-}\pi)$ for $B_* < \pi \le 1$ (see *Figure* 4.3 below). Note that $\pi \mapsto V_*(\pi)$ is C^1 on $[0,1]$ everywhere but at B_* where it is C^0 . To complete the proof it is enough to show that such defined map $\pi \mapsto V_*(\pi)$ equals the payoff defined in (4.2.8), and that τ_* defined in (4.2.36) is an optimal stopping time.

2. Since $\pi \mapsto V_*(\pi)$ is not C^1 only at one point at which it is C^0 , we can apply Itô's formula to $V_*(\pi_t)$. In exactly the same way as in (4.2.17) this gives:

$$V_*(\pi_t) \;=\; V_*(\pi) + \int_0^t (I\!\!L V_*)(\pi_{s-})\, ds + M_t \tag{4.2.40}$$

where $M = (M_t)_{t \ge 0}$ is a martingale given by:

$$M_t \;=\; \int_0^t \Big(V_*\big(\pi_{s-} + \Delta \pi_s\big) - V_*(\pi_{s-}) \Big)\, d\widehat{X}_s \tag{4.2.41}$$

and $\widehat{X}_t = X_t - \int_0^t E_\pi(\lambda \,|\, \mathcal{F}_{s-}^X)\, ds = X_t - \int_0^t (\lambda_1 \pi_{s-} + \lambda_0 (1{-}\pi_{s-}))\, ds$ is the so-called *innovation process* (see e.g. [49; Theorem 18.3]) which is a martingale with respect to $(\mathcal{F}_t^X)_{t \ge 0}$ and P_π whenever $\pi \in [0,1]$. Note in (4.2.40) that we may extend V_*' arbitrarily to B_* as the time spent by the process $(\pi_t)_{t \ge 0}$ at B_* is of Lebesgue measure zero.

3. Recall that $(I\!\!L V_*)(\pi) = -1$ for $\pi \in \big\langle A_*, B_* \big\rangle$, and note that due to the smooth fit (4.2.23) we also have $(I\!\!L V_*)(\pi) \ge -1$ for all $\pi \in [0,1] \setminus \big\langle A_*, B_* \big]$.

To verify this claim first note that $(I\!\!L V_*)(\pi) = 0$ for $\pi \in \big\langle 0, S^{-1}(A_*) \big\rangle \cup \big\langle B_*, 1 \big\rangle$, since $I\!\!L f \equiv 0$ if $f(\pi) = a\pi$ or $f(\pi) = b(1-\pi)$. Observe also that $(I\!\!L V_*)(S^{-1}(A_*)) = 0$ and $(I\!\!L V_*)(A_*) = -1$ both due to the smooth fit (4.2.23). Thus, it is enough to verify that $(I\!\!L V_*)(\pi) \ge -1$ for $\pi \in \big\langle S^{-1}(A_*), A_* \big\rangle$.

For this, consider the equation $I\!\!L V = -1$ on $\big\langle S^{-1}(A_*), A_* \big]$ upon imposing $V(\pi) = V(\pi; B_*)$ for $\pi \in \big\langle A_*, S(A_*) \big]$, and solve it under the initial condition $V(A_*) = V(A_*; B_*) + c$ where

$c \geq 0$. This generates a unique solution $\pi \mapsto V_c(\pi)$ on $\langle S^{-1}(A_*), A_*]$, and from (4.2.28) we read that $V_c(\pi) = V(\pi; B_*) + K_c(1-\pi)^{\gamma_1}/\pi^{\gamma_0}$ for $\pi \in \langle S^{-1}(A_*), A_*]$ where $K_c = c(A_*)^{\gamma_0}/(1-A_*)^{\gamma_1}$. (Observe that the curves $\pi \mapsto V_c(\pi)$ do not intersect on $\langle S^{-1}(A_*), A_*]$ for different c's.) Hence we see that there exists $c_0 > 0$ large enough such that for each $c > c_0$ the curve $\pi \mapsto V_c(\pi)$ lies strictly above the curve $\pi \mapsto a\pi$ on $\langle S^{-1}(A_*), A_*]$, and for each $c < c_0$ the two curves intersect. For $c \in [0, c_0\rangle$ let π_c denote the (closest) point (to A_*) at which $\pi \mapsto V_c(\pi)$ intersects $\pi \mapsto a\pi$ on $\langle S^{-1}(A_*), A_*]$. Then $\pi_0 = A_*$ and π_c decreases (continuously) in the direction of $S^{-1}(A_*)$ when c increases from 0 to c_0 . Observe that the points π_c are 'good' points since by $V_c(\pi_c) = a\pi_c = V_*(\pi_c)$ with $V_c'(\pi_c) > a = V_*'(\pi_c)$ and $V_c(S(\pi_c)) = V(S(\pi_c); B_*) = V_*(S(\pi_c))$ we see from (4.2.19) that $(I\!\!L V_*)(\pi_c) \geq (I\!\!L V_c)(\pi_c) = -1$. Thus, if we show that π_c reaches $S^{-1}(A_*)$ when $c \uparrow c_0$, then the proof of the claim will be complete. Therefore assume on the contrary that this is not the case. Then $V_{c_1}(S^{-1}(A_*)-) = aS^{-1}(A_*)$ for some $c_1 < c_0$, and $V_c(S^{-1}(A_*)-) > aS^{-1}(A_*)$ for all $c > c_1$. Thus by choosing $c > c_1$ close enough to c_1 , we see that a point $\widetilde{\pi}_c > S^{-1}(A_*)$ arbitrarily close to $S^{-1}(A_*)$ is obtained at which $V_c(\widetilde{\pi}_c) = a\widetilde{\pi}_c = V_*(\widetilde{\pi}_c)$ with $V_c'(\widetilde{\pi}_c) < a = V_*'(\widetilde{\pi}_c)$ and $V_c(S(\widetilde{\pi}_c)) = V(S(\widetilde{\pi}_c); B_*) = V_*(S(\widetilde{\pi}_c))$, from where we again see by (4.2.19) that $(I\!\!L V_*)(\widetilde{\pi}_c) \leq (I\!\!L V_c)(\widetilde{\pi}_c) = -1$. This however leads to a contradiction because $\pi \mapsto (I\!\!L V_*)(\pi)$ is continuous at $S^{-1}(A_*)$ (due to the smooth fit) and $(I\!\!L V_*)(S^{-1}(A_*)) = 0$ as already stated earlier. Thus, we have $(I\!\!L V_*)(\pi) \geq -1$ for all $\pi \in [0,1]$ (upon setting $V_*'(B_*) := 0$ for instance).

4. Recall further that $V_*(\pi) \leq g_{a,b}(\pi)$ for all $\pi \in [0,1]$. Moreover, since $\pi \mapsto V_*(\pi)$ is bounded, and $(X_t - \lambda_i\, t\,)_{t \geq 0}$ is a martingale under P_i for $i = 0, 1$, it is easily seen from (4.2.41) with (4.2.17) upon using the optional sampling theorem, that $E_\pi(M_\tau) = 0$ whenever τ is a stopping time of X such that $E_\pi(\tau) < \infty$. Thus, taking the expectation on both sides in (4.2.40), we obtain the inequality:

$$V_*(\pi) \;\leq\; E_\pi\big(\, \tau + g_{a,b}(\pi_\tau)\,\big) \tag{4.2.42}$$

for all such stopping times, and hence $V_*(\pi) \leq V(\pi)$ for all $\pi \in [0,1]$.

5. On the other hand, the stopping time τ_* from (4.2.36) clearly satisfies $V_*(\pi_{\tau_*}) = g_{a,b}(\pi_{\tau_*})$. Moreover, a direct analysis of τ_* based on (4.2.12)-(4.2.14) (see Remark 4.2.3 below), together with the fact that for a Poisson process $(N_t)_{t \geq 0}$ the exit time of the process $(N_t - \mu\, t)_{t \geq 0}$ from $[\widetilde{A}, \widetilde{B}]$ has a finite expectation for any real μ , shows that $E_\pi(\tau_*) < \infty$ for all $\pi \in [0,1]$. Taking then the expectation on both sides in (4.2.40), we get:

$$V_*(\pi) \;=\; E_\pi\big(\, \tau_* + g_{a,b}(\pi_{\tau_*})\big) \tag{4.2.43}$$

for all $\pi \in [0,1]$. This fact and the consequence of (4.2.42) stated above show that $V_* = V$, and that τ_* is an optimal stopping time. The proof of the first part is complete.

(II): Although, in principle, it is clear from our construction above that the second part of the theorem holds as well, we shall present a formal argument for completeness.

Suppose that the π-Bayes decision rule is not trivial. In other words, this means that $V(\pi) < g_{a,b}(\pi)$ for some $\pi \in \langle 0, 1\rangle$. Since $\pi \mapsto V(\pi)$ is concave, this implies that there are $0 < A_* < b/(a+b) < B_* < 1$ such that $\tau_* = \inf\,\{\, t > 0 \mid \pi_t \notin \langle A_*, B_*\rangle\,\}$ is optimal for the problems (4.2.8) and (4.2.5) respectively, with d_* from (4.2.9) in the latter case. Thus

$V(\pi) = E_\pi(\tau_* + g_{a,b}(\pi_{\tau_*}))$ for $\pi \in [0,1]$, and therefore by the general Markov processes theory, and due to the strong Markov property of $(\pi_t)_{t\geq 0}$, we know that $\pi \mapsto V(\pi)$ solves (4.2.20) and satisfies (4.2.21) and (4.2.22); a priori we do not know if the smooth fit condition (4.2.23) is satisfied. Nevertheless, these arguments show the existence of a solution to (4.2.20) on $\langle 0, B_*]$ which is $b(1-B_*)$ at B_* and which crosses $\pi \mapsto a\pi$ at (some) $A_* < b/(a+b)$. But then the same uniqueness argument used in Subsection 8 above (see Remark 4.2.2 below) shows that there must exist points $\widehat{A}_* \leq A_*$ and $\widehat{B}_* \geq B_*$ such that the solution $\pi \mapsto \widehat{V}(\pi; \widehat{B}_*)$ of (4.2.20) satisfying $\widehat{V}(\widehat{B}_*; \widehat{B}_*) = b(1-\widehat{B}_*)$ hits $\pi \mapsto a\pi$ smoothly at $\widehat{A}_*$. The first part of the proof above then shows that the stopping time $\widehat{\tau}_* = \inf\{\, t>0 \mid \pi_t \notin \langle \widehat{A}_*, \widehat{B}_* \rangle \,\}$ is optimal. As this stopping time is known to be P_π-a.s. pointwise the smallest possible optimal stopping time (see the proof of Theorem 4.3.1 below), this shows that τ_* cannot be optimal unless the smooth fit condition holds at A_* , that is, unless $\widehat{A}_* = A_*$ and $\widehat{B}_* = B_*$. In any case, however, this argument implies the existence of a non-trivial solution to the system (4.2.20)-(4.2.23), and since this fact is equivalent to (4.2.35) as shown above, we see that condition (4.2.35) cannot be violated.

Observe that we have actually proved that if the optimal stopping problem (4.2.8) has a non-trivial solution, then the principle of smooth fit holds at A_* . An alternative proof of the statement could be done by using Lemma 3 in [65; p.118]. The proof of the theorem is complete. $\square$

Remark 4.2.2

The following probabilistic argument can be given to show that the two curves $\pi \mapsto V(\pi, B')$ and $\pi \mapsto V(\pi, B'')$ from (4.2.32) do not intersect on $\langle 0, B']$ whenever $0 < B' < B'' \leq 1$.

Assume that the two curves do intersect at some $Z < B'$. Let $\pi \mapsto \alpha\pi + \beta$ denote the tangent of the map $V(\,\cdot\,; B')$ at Z . Define a map $\pi \mapsto g(\pi)$ by setting $g(\pi) = (\alpha\pi + \beta) \wedge b(1-\pi)$ for $\pi \in [0,1]$, and consider the optimal stopping problem (4.2.8) with g instead of $g_{a,b}$. Let $V = V(\pi)$ denote the value function. Consider also the map $\pi \mapsto V_*(\pi)$ defined by $V_*(\pi) = V(\pi; B')$ for $\pi \in [Z, B']$ and $V_*(\pi) = g(\pi)$ for $\pi \in [0,1] \setminus [Z, B']$. As $\pi \mapsto V_*(\pi)$ is C^0 at B' and C^1 at Z , then in exactly the same way as in Subsections 3-5 (Part I) of the proof above we find that $V_*(\pi) = V(\pi)$ for all $\pi \in [0,1]$. However, if we consider the stopping time $\sigma_* = \inf\{\, t>0 \mid \pi_t \notin \langle Z, B'' \rangle \,\}$, then it follows in the same way as in Subsection 5 (Part I) of the proof above that $V(\pi; B'') = E_\pi(\sigma_* + g(\pi_{\sigma_*}))$ for all $\pi \in [Z, B'']$. As $V(\pi; B'') < V_*(\pi)$ for $\pi \in \langle Z, B']$, this is a contradiction. Thus, the curves do not intersect.

Remark 4.2.3

1. Observe that the optimal decision rule (4.2.36)-(4.2.37) can be rewritten as follows:

$$\tau_* = \inf\left\{\, t \geq 0 \mid Z_t \notin \langle \widetilde{A}_*, \widetilde{B}_* \rangle \,\right\} \tag{4.2.44}$$

$$\begin{aligned} d_* &= 1 \quad (\text{accept } H_1)\ , \quad \text{if } Z_{\tau_*} \geq \widetilde{B}_* \\ &= 0 \quad (\text{accept } H_0)\ , \quad \text{if } Z_{\tau_*} = \widetilde{A}_* \end{aligned} \tag{4.2.45}$$

where we use the following notation:

$$Z_t = X_t - \mu t \tag{4.2.46}$$

(4.2.47) $$\widetilde{A}_* = \log\left(\frac{A_*}{1-A_*}\,\frac{1-\pi}{\pi}\right)\Big/\log\left(\frac{\lambda_1}{\lambda_0}\right)$$

(4.2.48) $$\widetilde{B}_* = \log\left(\frac{B_*}{1-B_*}\,\frac{1-\pi}{\pi}\right)\Big/\log\left(\frac{\lambda_1}{\lambda_0}\right)$$

(4.2.49) $$\mu = \left(\lambda_1-\lambda_0\right)\Big/\log\left(\frac{\lambda_1}{\lambda_0}\right) .$$

2. The representation (4.2.44)-(4.2.45) reveals the structure and applicability of the optimal decision rule in a clearer manner. The result proved above shows that the following sequential procedure is optimal: *While observing* X_t *, monitor* Z_t *, and stop the observation as soon as* Z_t *enters either* $\langle -\infty, \widetilde{A}_*]$ *or* $[\widetilde{B}_*, \infty\rangle$ *; in the first case conclude* $\lambda = \lambda_0$ *, in the second conclude* $\lambda = \lambda_1$ *.*

In this process the condition (4.2.35) must be satisfied, and the constants A_* and B_* should be determined as a unique solution of the system (4.2.38)-(4.2.39). This system can be successfully treated by means of standard numerical methods if one mimics our travel from B_* to A_* in the construction of our solution in Subsection 8 above. A pleasant fact is that only a few steps by (4.2.24) will be often needed to recapture A_* if one starts from B_* .

3. After we completed our work we observed that the same problem was treated by different methods in [59]. It is interesting to note that we could not find any later reference to that work. We also observed that the necessary and sufficient condition (4.2.35) of Theorem 4.2.1 is different from the condition $a\lambda_1+b(\lambda_0+\lambda_1)<b/a$ found in [59].

4.3 Solution of the variational problem

In the variational formulation of the problem it is assumed that the sequentially observed process $X=(X_t)_{t\geq 0}$ is a Poisson process with intensity λ_0 or λ_1 , and no probabilistic assumption is made about the outcome of λ_0 and λ_1 at time 0 . To formulate the problem we shall adopt the setting and notation from the previous section. Thus P_i is a probability measure on $(\Omega, \mathcal{F})$ under which $X=(X_t)_{t\geq 0}$ is a Poisson process with intensity λ_i for $i=0,1$.

1. Given the numbers α , $\beta > 0$ such that $\alpha+\beta<1$, let $\Delta(\alpha,\beta)$ denote the class of all decision rules (τ,d) satisfying:

(4.3.1) $$\alpha(d) \leq \alpha \quad \text{and} \quad \beta(d) \leq \beta$$

where $\alpha(d) = P_1(d=0)$ and $\beta(d) = P_0(d=1)$. The variational problem is then to find a decision rule $(\widehat{\tau}, \widehat{d}\,)$ in the class $\Delta(\alpha,\beta)$ such that:

(4.3.2) $$E_0(\widehat{\tau}) \leq E_0(\tau) \quad \text{and} \quad E_1(\widehat{\tau}) \leq E_1(\tau)$$

for any other decision rule (τ,d) from the class $\Delta(\alpha,\beta)$. Note that the main virtue of the requirement (4.3.2) is its simultaneous validity for both P_0 and P_1 .

This formulation of the problem is due to Wald [70]. In the papers [72] and [73] Wald and Wolfowitz proved the optimality of the SPRT in the case of i.i.d. observations and under special assumptions on the admissibility of (α,β) (see [72, 73, 1, 47] for more details and compare it with the admissability notion given below). In the paper [25] Dvoretzky, Kiefer and Wolfowitz considered the problem of optimality of the SPRT in the case of a continuous time and satisfied themselves with the remark that "a careful examination of the results of [72] and [73] shows that their conclusions in no way require that the processes be discrete in time" omitting any further detail and concentrating their attention to the problem of finding the error probabilities $\alpha(d)$ and $\beta(d)$ with expectations $E_0(\tau)$ and $E_1(\tau)$ for the given SPRT (τ, d) defined by "stopping boundaries" A and B in the cases of a Wiener and Poisson process. The explicit solution of the Bayesian problem in the case of a Wiener process was given in [64] (see also [65]). For the general problem of the minimax optimality of the SPRT for the case of a continuous time see [38].

Our main aim in this section is to show how the solution of the variational problem together with a precise description of all admissible pairs (α, β) can be obtained from the Bayesian solution in the previous section. The sequential procedure which leads to the optimal decision rule $(\widehat{\tau}, \widehat{d}\,)$ in this process is a SPRT which (as already mentioned earlier) was studied for the first time in [25]. We now describe a well-known procedure of passing from the Bayesian solution to the variational solution with some basic facts from [25] adapted to our aims.

2. Note that the explicit procedure of passing from the Bayesian solution to the variational solution presented in the next three steps is not confined to a Poissonian case but is also valid in greater generality (cf. [47]).

Step 1 (Construction): Given $\alpha\, , \beta > 0$ with $\alpha + \beta < 1$, find constants A and B satisfying $A < 0 < B$ such that the stopping time:

$$\widehat{\tau} = \inf \left\{\, t \geq 0 \mid Z_t \notin \left\langle A, B \right\rangle \,\right\} \tag{4.3.3}$$

satisfies the following identities:

$$P_1\Big(Z_{\hat{\tau}} = A \Big) = \alpha \tag{4.3.4}$$

$$P_0\Big(Z_{\hat{\tau}} \geq B \Big) = \beta \tag{4.3.5}$$

where $(Z_t)_{t \geq 0}$ is as in (4.2.46). Associate with $\widehat{\tau}$ the following decision function:

$$\begin{aligned} \widehat{d} &= 1 \quad (\text{ accept } H_1) \,, \quad \text{if} \quad Z_{\hat{\tau}} \geq B \\ &= 0 \quad (\text{ accept } H_0) \,, \quad \text{if} \quad Z_{\hat{\tau}} = A \,. \end{aligned} \tag{4.3.6}$$

We shall actually see below that not for all values α and β such A and B exist; a function $G : \left\langle 0, 1 \right\rangle \rightarrow \left\langle 0, 1 \right\rangle$ is displayed in (4.3.24) such that the solution (A, B) to (4.3.4)-(4.3.5) exists only for $\beta \in \left\langle 0, G(\alpha) \right\rangle$ if $\alpha \in \left\langle 0, 1 \right\rangle$. Such values α and β will be called *admissible*.

Step 2 (Embedding): Once A and B are found for admissible α and β , we may respectively identify them with $\widetilde{A}_*$ and $\widetilde{B}_*$ from (4.2.47) and (4.2.48). Then, for any $\widehat{\pi} \in \left\langle 0, 1 \right\rangle$ given and fixed, we can uniquely determine A_* and B_* satisfying $0 < A_* < B_* < 1$ such that (4.2.47) and (4.2.48) hold with $\pi = \widehat{\pi}$. Once A_* and B_* are given, we can choose $a > 0$ and $b > 0$ in the Bayesian problem (4.2.4)+(4.2.5) such that the optimal stopping time in (4.2.8) is exactly the exit time τ_* of $(\pi_t)_{t \geq 0}$ from $\left\langle A_*, B_* \right\rangle$ as given in (4.2.36). Observe that this is possible to achieve since the optimal A_* and B_* range through all $\left\langle 0, 1 \right\rangle$ when a and b satisfying

(4.2.35) range through $\langle 0, \infty \rangle$. (For this, let any $B_* \in \langle 0,1 \rangle$ be given and fixed, and choose $\widetilde{a} > 0$ and $\widetilde{b} > 0$ such that $B_* = \widetilde{b}/(\widetilde{a}+\widetilde{b})$ with $\lambda_1 - \lambda_0 = 1/\widetilde{a} + 1/\widetilde{b}$. Then consider the solution $V(\,\cdot\,; B_*) := V_b(\,\cdot\,; B_*)$ of (4.2.20) on $\langle 0, B_* \rangle$ upon imposing $V_b(\pi; B_*) = b(1-\pi)$ for $\pi \in [B_*, S(B_*)]$ where $b \geq \widetilde{b}$. To each such a solution there corresponds $a > 0$ such that $\pi \mapsto a\pi$ hits $\pi \mapsto V_b(\pi; B_*)$ smoothly at some $A = A(b)$. When b increases from $\widetilde{b}$ to ∞ , then $A(b)$ decreases from B_* to zero. This is easily verified by a simple comparison argument upon noting that $\pi \mapsto V_b(\pi; B_*)$ stays strictly above $\pi \mapsto V(\pi; B_*) + V_b(B_*; B_*)$ on $\langle 0, B_* \rangle$ (recall the idea used in Remark 4.2.3 above). As each $A(b)$ obtained (in the pair with B_*) is optimal (recall the arguments used in Subsections 3-5 (Part I) of the proof of Theorem 4.2.1), the proof of the claim is complete.)

Step 3 (Verification): Consider the process $(\widehat{\pi}_t)_{t \geq 0}$ defined by (4.2.12)+(4.2.14) with $\pi = \widehat{\pi}$, and denote by $(\widehat{\tau}_*, \widehat{d}_*)$ the optimal decision rule (4.2.36)-(4.2.37) associated with it. From our construction above note that $\widehat{\tau}$ from (4.3.3) actually coincides with $\widehat{\tau}_*$, as well as that $(\widehat{\pi}_{\widehat{\tau}_*} = A_*) = (Z_{\hat{\tau}} = A)$ and $(\widehat{\pi}_{\widehat{\tau}_*} \geq B_*) = (Z_{\hat{\tau}} \geq B)$. Thus (4.3.4) and (4.3.5) show that:

$$P_1\Big(\widehat{d}_* = 0\Big) = \alpha \tag{4.3.7}$$

$$P_0\Big(\widehat{d}_* = 1\Big) = \beta \tag{4.3.8}$$

for the admissible α and β . If now any decision rule (τ, d) from $\Delta(\alpha, \beta)$ is given, then either $P_1(d=0) = \alpha$ and $P_0(d=1) = \beta$, or at least one strict inequality holds. In both cases, however, from (4.2.4)-(4.2.6) and (4.3.7)+(4.3.8) we easily see that $E_{\widehat{\pi}}(\widehat{\tau}_*) \leq E_{\widehat{\pi}}(\tau)$, since otherwise $\widehat{\tau}_*$ would not be optimal. Since $\widehat{\tau}_* = \widehat{\tau}$, it follows $E_{\widehat{\pi}}(\widehat{\tau}) \leq E_{\widehat{\pi}}(\tau)$, and letting $\widehat{\pi}$ first to 0 and then to 1 , we obtain (4.3.2) in the case when $E_0(\tau) < \infty$ and $E_1(\tau) < \infty$. If either $E_0(\tau)$ or $E_1(\tau)$ equals ∞ , then (4.3.2) follows by the same argument after a simple truncation (e.g. if $E_0(\tau) < \infty$ but $E_1(\tau) = \infty$ choose $n \geq 1$ such that $P_0(\tau > n) \leq \varepsilon$, apply the same argument to $\tau_n := \tau \wedge n$ and $d_n := d\, 1_{\{\tau \leq n\}} + 1_{\{\tau > n\}}$, and let ε go to zero in the end.) This solves the variational problem posed above for all admissible α and β .

3. The preceding arguments also show:

(4.3.9) If either $P_1(d=0) < \alpha$ or $P_0(d=1) < \beta$ for some $(\tau, d) \in \Delta(\alpha, \beta)$ with admissible α and β , then at least one strict inequality in (4.3.2) holds.

Moreover, since $\widehat{\tau}_*$ is known to be $P_{\widehat{\pi}}$-a.s. the smallest possible optimal stopping time (see the proof of Theorem 4.3.1 below), from the arguments above we also get:

(4.3.10) If $P_1(d=0) = \alpha$ and $P_0(d=1) = \beta$ for some $(\tau, d) \in \Delta(\alpha, \beta)$ with admissible α and β , and both equalities in (4.3.2) hold, then $\tau = \widehat{\tau}$ P_0-a.s. and P_1-a.s.

The property (4.3.10) characterises $\widehat{\tau}$ as *a unique stopping time of the decision rule with maximal admissible error probabilities having both* P_0 *and* P_1 *expectation at minimum.*

4. It remains to determine admissible α and β in (4.3.4) and (4.3.5) above. For this, consider $\widehat{\tau}$ defined in (4.3.3) for some $A < 0 < B$, and note from (4.2.14) that $\varphi_t = \exp\big(Z_t \log(\lambda_1/\lambda_0)\big)$. By means of (4.2.10) we find:

$$(4.3.11) \qquad P_1\Big\{Z_{\hat\tau}=A\Big\} = P_1\left\{\varphi_{\hat\tau}=\exp\Big(A\log\Big(\frac{\lambda_1}{\lambda_0}\Big)\Big)\right\}$$

$$= \exp\Big(A\log\Big(\frac{\lambda_1}{\lambda_0}\Big)\Big)P_0\Big\{Z_{\hat\tau}=A\Big\} = \exp\Big(A\log\Big(\frac{\lambda_1}{\lambda_0}\Big)\Big)\Big(1-P_0\Big\{Z_{\hat\tau}\ge B\Big\}\Big)\ .$$

Using (4.3.4)-(4.3.5), from (4.3.11) we see that:

$$(4.3.12) \qquad A = \log\Big(\frac{\alpha}{1-\beta}\Big)\Big/\log\Big(\frac{\lambda_1}{\lambda_0}\Big)\ .$$

To determine B , let P_0^z be a probability measure under which $(X_t)_{t\ge0}$ is a Poisson process with intensity λ_0 and $(Z_t)_{t\ge0}$ starts at z . It is easily seen that the infinitesimal operator of $(Z_t)_{t\ge0}$ under $(P_0^z)_{z\in I\!R}$ acts like:

$$(4.3.13) \qquad (I\!L_0 f)(z) = -\mu f'(z)+\lambda_0\Big(f(z+1)-f(z)\Big)\ .$$

In view of (4.3.5), introduce the function:

$$(4.3.14) \qquad u(z) = P_0^z\Big(Z_{\hat\tau}\ge B\Big)\ .$$

Strong Markov arguments then show that $z\mapsto u(z)$ solves the following system:

$$(4.3.15) \qquad (I\!L_0 u)(z) = 0 \quad \text{if} \quad z\in\langle A,B\rangle\setminus\{B-1\}$$

$$(4.3.16) \qquad u(A) = 0$$

$$(4.3.17) \qquad u(z) = 1 \quad \text{if} \quad z\ge B\ .$$

The solution of this system is given in (4.15) of [25]. To display it, introduce the function:

$$(4.3.18) \qquad F(x;B) = \sum_{k=0}^{\delta(x,B)}\frac{(-1)^k}{k!}\Big(\big(B-x-k\big)\,\rho\,e^{-\rho}\Big)^k$$

for $x\le B$, where we denote:

$$(4.3.19) \qquad \delta(x,B) = -[x-B+1]$$

$$(4.3.20) \qquad \rho = \log\Big(\frac{\lambda_1}{\lambda_0}\Big)\Big/\Big(\frac{\lambda_1}{\lambda_0}-1\Big)\ .$$

Setting $J_n=[B-n-1,B-n\rangle$ for $n\ge0$, observe that $\delta(x,B)=n$ if and only if $x\in J_n$.

It is then easily verified that the solution of the system (4.3.15)-(4.3.17) is given by:

$$(4.3.21) \qquad u(z) = 1-e^{-\rho(z-A)}\,\frac{F(z;B)}{F(A;B)}$$

for $A\le z<B$. Note that $z\mapsto u(z)$ is C^1 everywhere in $\langle A,B\rangle$ but at $B-1$ where it is only C^0 ; note also that $u(A+)=u(A)=0$, but $u(B-)<u(B)=1$ (see *Figure* 4.4 below).

Going back to (4.3.5), and using (4.3.21), we see that:

$$P_0\Big(Z_{\hat\tau} \geq B\Big) = 1 - e^{\rho A}\,\frac{F(0;B)}{F(A;B)}\ . \tag{4.3.22}$$

Letting $B \downarrow 0$ in (4.3.22), and using the fact that the expression (4.3.22) is continuous in B and decreases to 0 as $B \uparrow \infty$, we clearly obtain a necessary and sufficient condition on β to satisfy (4.3.5), once $A = A(\alpha,\beta)$ is fixed through (4.3.12); as $F(0;0) = 1$, this condition reads:

$$\beta < 1 - \frac{e^{\rho A(\alpha,\beta)}}{F\big(A(\alpha,\beta);0\big)}\ . \tag{4.3.23}$$

Note, however, if β increases, then the function on the right-hand side in (4.3.23) decreases, and thus there exists a unique $\beta_* = \beta_*(\alpha) > 0$ at which equality in (4.3.23) is attained. (This value can easily be computed by means of standard numerical methods.) Setting:

$$G(\alpha) = 1 - \frac{e^{\rho A(\alpha,\beta_*(\alpha))}}{F\big(A(\alpha,\beta_*(\alpha));0\big)} \tag{4.3.24}$$

we see that admissible α and β are characterised by $0 < \beta < G(\alpha)$ (see *Figure* 4.5 below). In this case A is given by (4.3.12), and B is uniquely determined from the equation:

$$F(0;B) - (1-\beta)\,F(A;B)\,e^{-\rho A} = 0\ . \tag{4.3.25}$$

The set of all admissible α and β will be denoted by $\mathcal{A}$. Thus, we have:

$$\mathcal{A} = \big\{\,(\alpha,\beta) \mid 0 < \alpha < 1\,,\ 0 < \beta < G(\alpha)\,\big\}\ . \tag{4.3.26}$$

5. The preceding considerations may be summarised as follows (see also Remark 4.3.2 below).

Theorem 4.3.1

In the problem (4.3.1)-(4.3.2) of testing two simple hypotheses (4.2.2)-(4.2.3) based upon sequential observations of the Poisson process $X = (X_t)_{t\geq 0}$ *under* P_0 *or* P_1 *, there exists a unique decision rule* $(\widehat\tau,\widehat d\,) \in \Delta(\alpha,\beta)$ *satisfying (4.3.2) for any other decision rule* $(\tau,d) \in \Delta(\alpha,\beta)$ *whenever* $(\alpha,\beta) \in \mathcal{A}$ *. The decision rule* $(\widehat\tau,\widehat d\,)$ *is explicitly given by (4.3.3)+(4.3.6) with* A *in (4.3.12) and* B *from (4.3.25), it satisfies (4.3.9), and is characterised by (4.3.10).*

Proof. It only remains to prove (4.3.10). For this, in the notation used above, assume that τ is a stopping time of X satisfying the hypotheses of (4.3.10). Then clearly τ is an optimal stopping time in (4.2.8) for $\pi = \widehat\pi$ with a and b as in Step 2 above.

Recall that $V_*(\pi) \leq g_{a,b}(\pi)$ for all π, and observe that $\widehat\tau$ can be written as:

$$\widehat\tau = \inf\big\{\,t \geq 0 \mid V_*(\widehat\pi_t) \geq g_{a,b}(\widehat\pi_t)\,\big\} \tag{4.3.27}$$

where $\pi \mapsto V_*(\pi)$ is the payoff (4.2.8) appearing in the proof of Theorem 4.2.1. Supposing now that $P_{\hat\pi}(\tau < \widehat\tau) > 0$, we easily find by (4.3.27) that:

$$E_{\hat\pi}\Big(\,\tau + g_{a,b}\big(\widehat\pi_\tau\big)\,\Big) > E_{\hat\pi}\Big(\,\tau + V_*\big(\widehat\pi_\tau\big)\,\Big)\ . \tag{4.3.28}$$

On the other hand, it is clear from (4.2.40) with $LV_* \geq -1$ that $(t + V_*(\widehat{\pi}_t))_{t \geq 0}$ is a submartingale. Thus by the optional sampling theorem it follows that:

$$E_{\hat{\pi}}\Big(\tau + V_*(\widehat{\pi}_\tau) \Big) \geq V_*(\widehat{\pi}) \ . \tag{4.3.29}$$

However, from (4.3.28) and (4.3.29) we see that τ cannot be optimal, and thus we must have $P_{\hat{\pi}}(\tau \geq \widehat{\tau}) = 1$. Moreover, since it follows from our assumption that $E_{\hat{\pi}}(\tau) = E_{\hat{\pi}}(\widehat{\tau})$, this implies that $\tau = \widehat{\tau}$ $P_{\hat{\pi}}$-a.s. Finally, as $P_i << P_{\hat{\pi}}$ for $i = 0, 1$, we see that $\tau = \widehat{\tau}$ both P_0-a.s. and P_1-a.s. The proof of the theorem is complete. □

Observe that the sequential procedure of the optimal decision rule $(\widehat{\tau}, \widehat{d})$ from Theorem 4.3.1 is precisely the SPRT. The explicit formulas for $E_0(\widehat{\tau})$ and $E_1(\widehat{\tau})$ are given in (4.22) of [25].

Remark 4.3.2

If $(\alpha, \beta) \notin \mathcal{A}$, that is, if $\beta \geq G(\alpha)$ for some α , $\beta > 0$ such that $\alpha + \beta < 1$, then no decision rule given by the SPRT-form (4.3.3)+(4.3.6) can solve the variational problem (4.3.1)-(4.3.2).

To see this, let such $(\alpha, \beta^*) \notin \mathcal{A}$ be given, and let (τ, d) be a decision rule satisfying (4.3.3)+(4.3.6) for some $A < 0 < B$. Denote $\beta = P_0(Z_\tau \geq B)$ and choose α to satisfy (4.3.12). Then $\beta < G(\alpha) \leq \beta^*$ by definition of the map G . Given $\beta' \in \big\langle \beta, G(\alpha) \big\rangle$, let B' be taken to satisfy (4.3.5) with β' , and let α' be determined from (4.3.12) with β' so that A remains unchanged. Clearly $0 < B' < B$ and $0 < \alpha' < \alpha$, and (4.3.4) holds with A and α' respectively. But then (τ', d') satisfying (4.3.3)+(4.3.6) with $A < 0 < B'$ still belongs to $\Delta(\alpha, \beta^*)$, while clearly $\tau' < \tau$ both under P_0 and P_1 . This shows that (τ, d) does not solve the variational problem.

The preceding argument shows that the admissible class $\mathcal{A}$ from (4.3.26) is exactly the class of all error probabilities (α, β) for which the SPRT is optimal. A pleasant fact is that $\mathcal{A}$ always contains a neighborhood around $(0, 0)$ in $[0, 1] \times [0, 1]$, which is the most interesting case from the point of view of statistical applications.

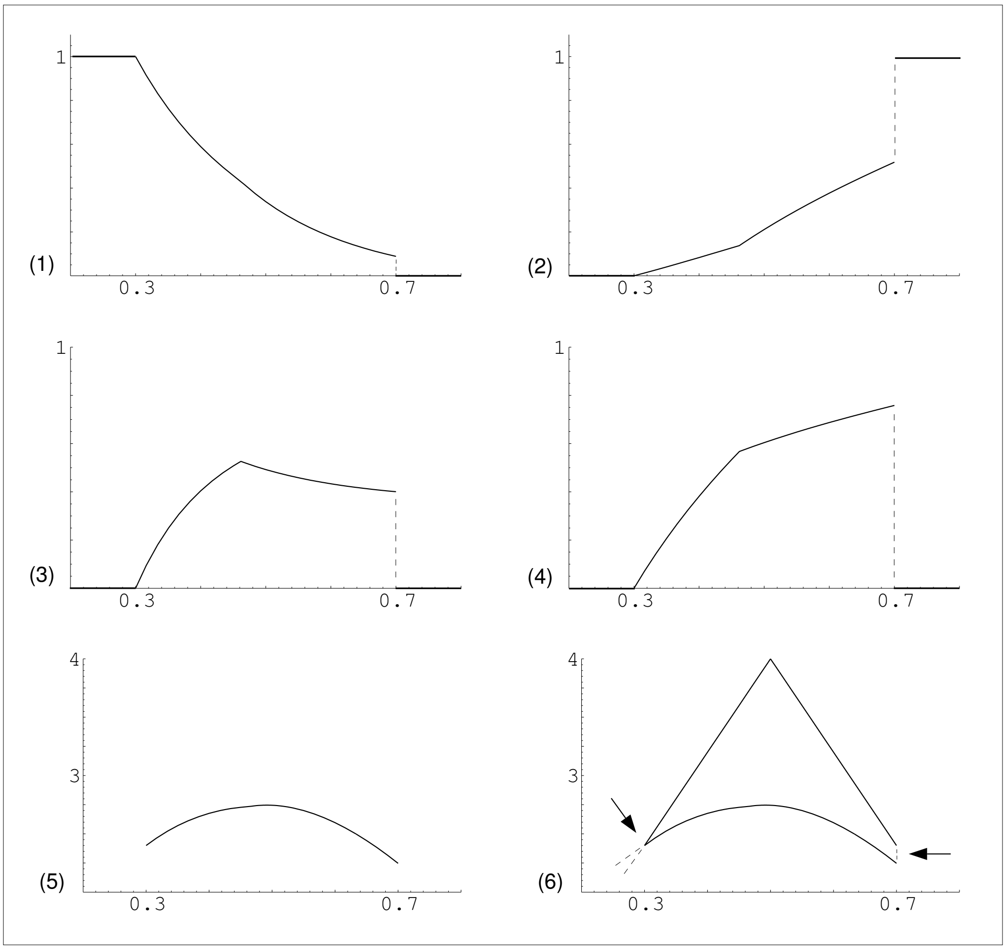


Figure 4.1. In view of the problem (4.2.8) and its decomposition via (4.2.4) and (4.2.6) with (4.2.1), we consider $\tau = \inf\{\, t \ge 0 \mid \pi_t \notin \langle A, B \rangle \}$ for $(\pi_t)_{t \ge 0}$ from (4.2.7)+(4.2.12)+(4.2.14) with $\pi \in \langle A, B \rangle$ given and fixed, so that $\pi_0 = \pi$ under P_0 and P_1 ; the computer drawings above show the following functions respectively: (1) $\pi \mapsto P_1(\pi_\tau = A)$; (2) $\pi \mapsto P_0(\pi_\tau \ge B)$; (3) $\pi \mapsto E_1(\tau)$; (4) $\pi \mapsto E_0(\tau)$; (5) $\pi \mapsto \pi E_1(\tau) + (1-\pi) E_0(\tau) + a\pi P_1(\pi_\tau = A) + b(1-\pi) P_0(\pi_\tau \ge B) = E_\pi(\tau + g_{a,b}(\pi_\tau))$; (6) $\pi \mapsto E_\pi(\tau + g_{a,b}(\pi_\tau))$ and $\pi \mapsto g_{a,b}(\pi)$, where $A = 0.3$, $B = 0.7$, $\lambda_0 = 1$, $\lambda_1 = e$ and $a = b = 8$. Functions (1)-(4) are found by solving systems analogous to the system (4.3.15)-(4.3.17); their discontinuity at B should be noted, as well as the discontinuity of their first derivative at $B_1 = 0.46\ldots$ from (4.2.25); observe that the function (5) is a superposition of functions (1)-(4), and thus the same discontinuities carry over to the function (5), unless something special occurs. The crucial fact to be observed is that if the function (5) is to be the payoff (4.2.8), and thus extended by the gain function $\pi \mapsto g_{a,b}(\pi)$ outside $\langle A, B \rangle$, then such an extension would generally be discontinuous at B and have a discontinuous first derivative at A ; this is depicted in the final picture (6). It is a matter of fact that the optimal A_* and B_* are to be chosen in such a way that both of these discontinuities disappear; these are the *principles of continuous and smooth fit* respectively. Observe that in this case the discontinuity of the first derivative of (5) also disappears at B_1 , and the extension obtained is C^1 everywhere but at B_* where it is only C^0 generally (see *Figure* 4.3 below).

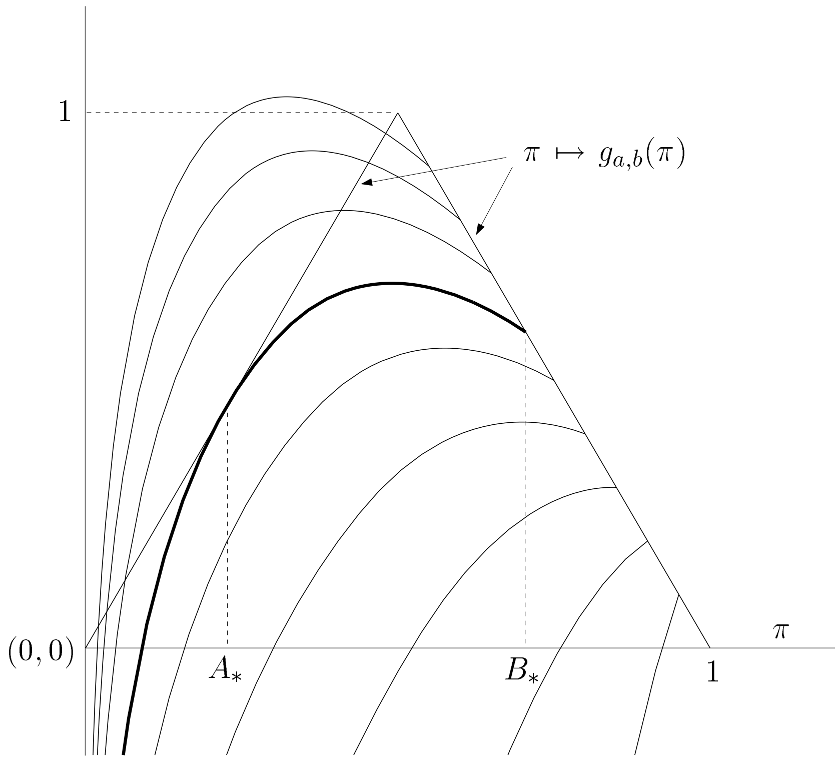


Figure 4.2. A computer drawing of "continuous fit" solutions $\pi \mapsto V(\pi; B)$ of (4.2.20), satisfying (4.2.21) on $[B, S(B)]$ and the second part of (4.2.22) at B, for different B in $\langle b/(a+b), 1 \rangle$; in this particular case we took $B = 0.95, 0.80, 0.75, \dots, 0.55$, with $\lambda_0 = 1$, $\lambda_1 = 5$ and $a = b = 2$. The unique B_* is obtained through the requirement that the map $\pi \mapsto V(\pi; B_*)$ hits "smoothly" the gain function $\pi \mapsto g_{a,b}(\pi)$ at A_*; as shown above, this happens for $A_* = 0.22\dots$ and $B_* = 0.70\dots$; such obtained A_* and B_* are a unique solution of the system (4.2.38)-(4.2.39). The solution $\pi \mapsto V(\pi; B_*)$ leads to the explicit form of the payoff (4.2.8) as shown in *Figure* 4.3 below.

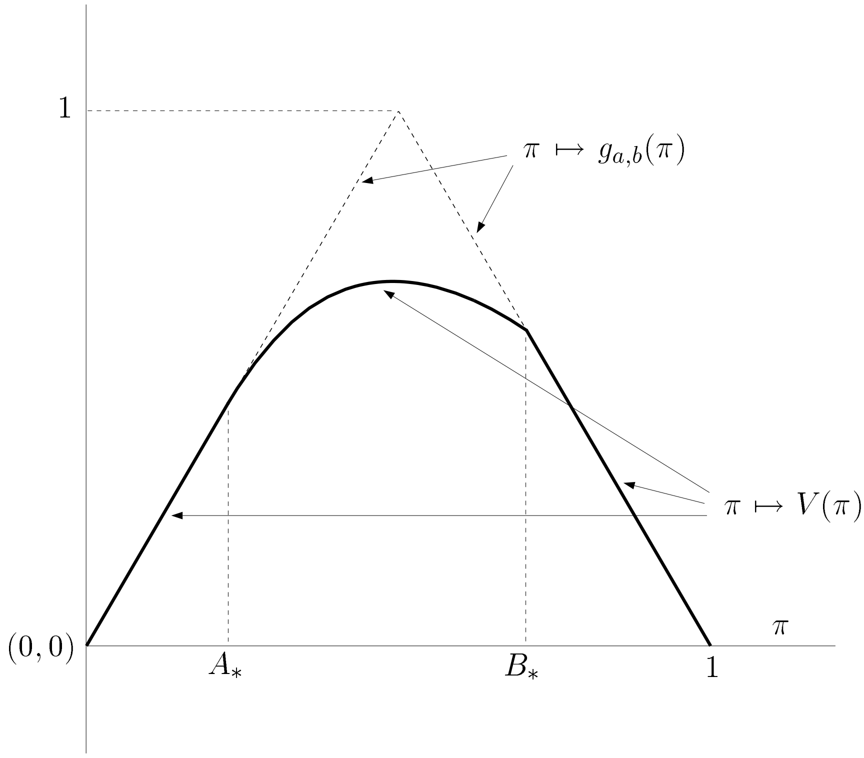


Figure 4.3. A computer drawing of the payoff (4.2.8) in the case $\lambda_0 = 1$, $\lambda_1 = 5$ and $a = b = 2$ as indicated in *Figure* 4.2 above. The interval $\langle A_*, B_* \rangle$ is the region of continued observation of the process $(\pi_t)_{t \geq 0}$, while its complement in $[0,1]$ is the stopping region. Thus, as indicated in (4.2.36), the observation should be stopped as soon as the process $(\pi_t)_{t \geq 0}$ enters $[0,1] \setminus \langle A_*, B_* \rangle$, and this stopping time is optimal in the problem (4.2.8). The optimal decision function is then given by (4.2.37).

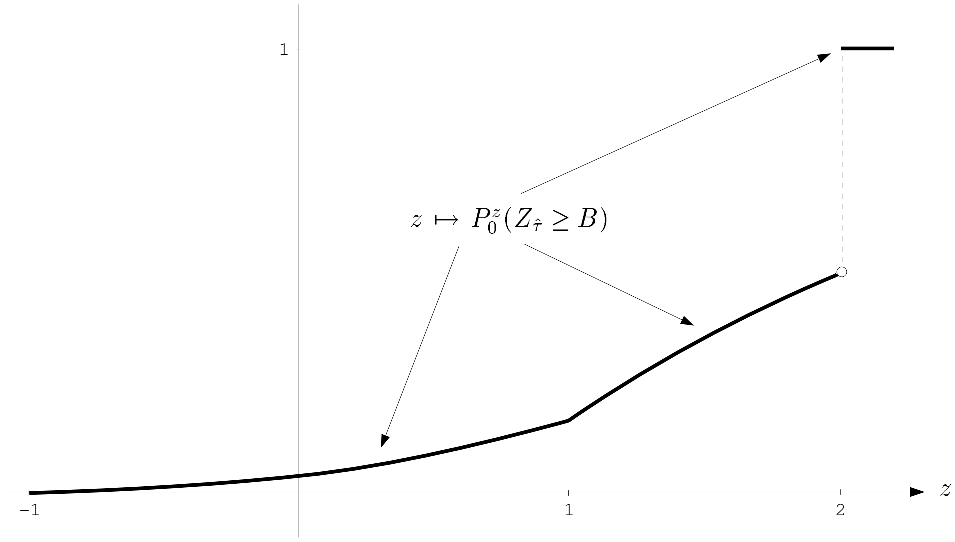


Figure 4.4. A computer drawing of the map $u(z) = P_0^z(Z_{\hat{\tau}} \ge B)$ from (4.3.14) in the case $A = -1$, $B = 2$ and $\lambda_0 = 0.5$. This map is a unique solution of the system (4.3.15)-(4.3.17). Its discontinuity at B should be noted, as well as the discontinuity of its first derivative at $B-1$. Observe also that $u(A+) = u(A) = 0$. The case of general A, B and λ_0 looks very much the same.

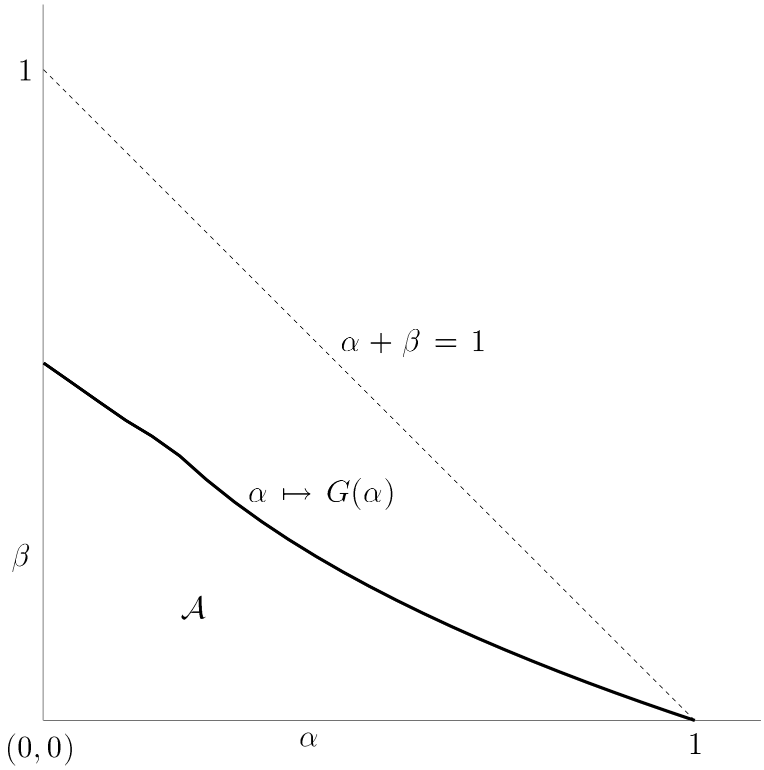


Figure 4.5. A computer drawing of the map $\alpha \mapsto G(\alpha)$ from (4.3.24) in the case $\lambda_0 = 1$ and $\lambda_1 = 3$. The area $\mathcal{A}$ which lies below the graph of G determines the set of all admissible α and β. The case of general λ_0 and λ_1 looks very much the same; it can also be shown that $G(0+)$ decreases if the difference $\lambda_1 - \lambda_0$ increases, as well as that $G(0+)$ increases if both λ_0 and λ_1 increase so that the difference $\lambda_1 - \lambda_0$ remains constant; in all cases $G(1-) = 0$. It may seem somewhat surprising that $G(0+) < 1$; observe, however, this is in agreement with the fact that $(Z_t)_{t \geq 0}$ from (4.2.46) is a supermartingale under P_0. (A little peak on the graph, at $\hat{\alpha} = 0.19\dots$ and $\hat{\beta} = 0.42\dots$ in this particular case, corresponds to the disturbance when A from (4.3.12) passes through -1 while $B = 0+$; it is caused by a discontinuity of the first derivative of the map from (4.3.22) at $B - 1$ (see *Figure* 4.4 above).)

5. Quickest detection problems

The Poisson disorder problem is less formally stated as follows. Suppose that at time $t = 0$ we begin observing a trajectory of the Poisson process $X = (X_t)_{t \geq 0}$ whose intensity changes from λ_0 to λ_1 at some random (unknown) time θ which is assumed to take value 0 with probability π , and is exponentially distributed with parameter λ given that $\theta > 0$. Based upon the information which is continuously updated through our observation of the trajectory of X , our problem is to terminate the observation (and declare the alarm) at a time τ_* which is as close as possible to θ (measured by a cost function with parameter $c > 0$ specified below).

5.1 Introduction

The problem above was first studied in [31] where a solution has been found in the case when $\lambda + c \geq \lambda_1 > \lambda_0$. This result has been extended in [21] to the case when $\lambda + c \geq \lambda_1 - \lambda_0 > 0$. Many other authors have also studied the problem from a different standpoint (see e.g. [51]). The main purpose of the present chapter is to describe the structure of the solution in the general case.

The Wiener process version of the disorder problem (where the drift changes) appeared earlier (see [64]) and is now well-understood (we refer to [65; p.208] for historical comments and references). The method of proof consists of reducing the initial (optimal stopping) problem to a free-boundary differential problem which can be solved explicitly. The principle of smooth fit plays a key role in this context.

In this chapter we adopt the same methodology as in the Wiener process case. A discontinuous character of the observed (Poisson) process in the present case, however, forces us to deal with a differential-difference equation forming a free-boundary problem which is more delicate. This in turn leads to a new effect of the breakdown of the smooth fit principle (and its replacement by the principle of continuous fit), and the key issue in the solution is to understand and specify when exactly this happens. This can be done, on one hand, in terms of the a posteriori probability process (i.e. its jump structure and sample path behaviour), and on the other hand, in terms of a singularity point of the equation from the free-boundary problem. Moreover, it turns out that the existence of such a singularity point makes explicit computations feasible.

The facts on the principles of continuous and smooth fit found here complement and further extend our findings in [P3]. Problems of detecting the arrival of 'disorder' are of central importance in quality control and have also found notable industrial and other applications.

5.2 The Poisson disorder problem

1. The Poisson disorder problem can be formally stated as follows. Let $N^{\lambda_0} = (N_t^{\lambda_0})_{t \geq 0}$, $N^{\lambda_1} = (N_t^{\lambda_1})_{t \geq 0}$ and $L = (L_t)_{t \geq 0}$ be three independent stochastic processes defined on a probability space $(\Omega, \mathcal{F}, P_\pi)$ with $\pi \in [0,1]$ such that:

(5.2.1) N^{λ_0} is a Poisson process with intensity $\lambda_0 > 0$;

(5.2.2) N^{λ_1} is a Poisson process with intensity $\lambda_1 > 0$;

(5.2.3) L is a continuous Markov chain with two states λ_0 and λ_1 , initial distribution $[1-\pi; \pi]$, and transition-probability matrix $[\, e^{-\lambda t} , 1 - e^{-\lambda t} \,; 0 \,, 1 \,]$ for $t > 0$ where $\lambda > 0$.

Thus $P_\pi(L_0 = \lambda_1) = 1 - P_\pi(L_0 = \lambda_0) = \pi$, and given that $L_0 = \lambda_0$, there is a single passage of L from λ_0 to λ_1 at a random time $\theta > 0$ satisfying $P_\pi(\theta > t) = e^{-\lambda t}$ for all $t > 0$.

The process $X = (X_t)_{t \geq 0}$ observed is given by:

$$X_t = \int_0^t I(L_{s-} = \lambda_0)\, dN_s^{\lambda_0} + \int_0^t I(L_{s-} = \lambda_1)\, dN_s^{\lambda_1} \tag{5.2.4}$$

and we set $\mathcal{F}_t^X = \sigma(X_s \,|\, 0 \leq s \leq t)$ for $t \geq 0$. Denoting $\theta = \inf\{\, t \geq 0 \,|\, L_t = \lambda_1 \,\}$ we see that $P_\pi(\theta = 0) = \pi$ and $P_\pi(\theta > t \,|\, \theta > 0) = e^{-\lambda t}$ for all $t > 0$. It is assumed that *the time* θ *of 'disorder'* is unknown (i.e. it cannot be observed directly).

The *Poisson disorder problem* seeks to find a stopping time τ_* of X that is 'as close as possible' to θ as a solution of the following optimal stopping problem:

$$V(\pi) = \inf_\tau \Big(P_\pi(\tau < \theta) + c E_\pi(\tau - \theta)^+ \Big) \tag{5.2.5}$$

where $P_\pi(\tau < \theta)$ is interpreted as the probability of a '*false alarm*', $E_\pi(\tau - \theta)^+$ is interpreted as the '*average delay*' in detecting the occurrence of 'disorder' correctly, $c > 0$ is a given constant, and the infimum in (5.2.5) is taken over all stopping times τ of X . [A stopping time of X means a stopping time with respect to the natural filtration $(\mathcal{F}_t^X)_{t \geq 0}$ generated by X . The same terminology will be used for other processes in the sequel as well.]

2. Introducing the *a posteriori probability process*:

$$\pi_t = P_\pi(\theta \leq t \mid \mathcal{F}_t^X) \tag{5.2.6}$$

for $t \geq 0$, it is easily seen that $P_\pi(\tau < \theta) = E_\pi(1 - \pi_\tau)$ and $E_\pi(\tau - \theta)^+ = E_\pi\big(\int_0^\tau \pi_t\, dt \big)$ for all stopping times τ of X , so that (5.2.5) can be rewritten as follows:

$$V(\pi) = \inf_\tau E_\pi \Big((1 - \pi_\tau) + c \int_0^\tau \pi_t\, dt \Big) \tag{5.2.7}$$

where the infimum is taken over all stopping times τ of $(\pi_t)_{t \geq 0}$ (as shown following (5.2.11) below).

Defining the *likelihood ratio process*:

$$\varphi_t = \frac{\pi_t}{1 - \pi_t} \tag{5.2.8}$$

it is possible to verify by standard means that the following explicit expression is valid:

$$\varphi_t = e^{\lambda t} e^{X_t \log(\lambda_1/\lambda_0) - (\lambda_1 - \lambda_0)t} \Big(\varphi_0 + \lambda \int_0^t e^{-\lambda s} e^{-X_s \log(\lambda_1/\lambda_0) + (\lambda_1 - \lambda_0)s}\, ds \Big) \tag{5.2.9}$$

for $t \geq 0$. Hence by Itô's formula (see e.g. [41]) one finds that the processes $(\varphi_t)_{t \geq 0}$ and $(\pi_t)_{t \geq 0}$ solve the following stochastic equations respectively:

$$d\varphi_t = \lambda(1 + \varphi_t)\, dt + \Big(\frac{\lambda_1}{\lambda_0} - 1 \Big) \varphi_{t-}\, d\big(X_t - \lambda_0 t \big) \tag{5.2.10}$$

$$d\pi_t = \lambda(1-\pi_t)\,dt + \frac{(\lambda_1-\lambda_0)\,\pi_{t-}(1-\pi_{t-})}{\lambda_1\pi_{t-} + \lambda_0\,(1-\pi_{t-})}\left(dX_t - \Big(\lambda_1\pi_{t-} + \lambda_0\,(1-\pi_{t-})\Big)dt\right) \tag{5.2.11}$$

(cf. [31; p.713] or [49; p.307]). It follows that $(\varphi_t)_{t\ge0}$ and $(\pi_t)_{t\ge0}$ are time-homogeneous (strong) Markov processes under P_π with respect to the natural filtrations which clearly coincide with $(\mathcal{F}_t^X)_{t\ge0}$ respectively. Thus, the infimum in (5.2.7) may indeed be viewed as taken over all stopping times τ of $(\pi_t)_{t\ge0}$, and the optimal stopping problem (5.2.7) falls into the class of optimal stopping problems for Markov processes. We thus proceed by finding the infinitesimal operator of the Markov process $(\pi_t)_{t\ge0}$.

By Itô's formula, upon making use of the fact easily verified (see (5.2.14) below) that the *innovation process* $\widehat{X}_t = X_t - \int_0^t E_\pi(L_s \mid \mathcal{F}_{s-}^X)\,ds = X_t - \int_0^t(\lambda_1\pi_{s-} + \lambda_0(1-\pi_{s-}))\,ds$ is a martingale under P_π with respect to $(\mathcal{F}_t^X)_{t\ge0}$, it follows from (5.2.11) that the infinitesimal operator of $(\pi_t)_{t\ge0}$ acts on $f \in C^1[0,1]$ according to the following rule:

$$\begin{aligned}(\mathbb{L}f)(\pi) = &\Big(\lambda-(\lambda_1-\lambda_0)\,\pi\Big)\,(1-\pi)\,f'(\pi)\\ &+\Big(\lambda_1\,\pi + \lambda_0\,(1-\pi)\Big)\left(f\left(\frac{\lambda_1\pi}{\lambda_1\pi+\lambda_0\,(1-\pi)}\right) - f(\pi)\right).\end{aligned} \tag{5.2.12}$$

It may be noted that the equations (5.2.10)-(5.2.12) for $\lambda = 0$ reduce to the analogous equations in [P3].

3. We may assume that for each $r \ge 0$ a probability measure Q_r is defined on $(\Omega, \mathcal{F})$ such that $Q_r(\theta = r) = 1$. Thus, under Q_r the observed process $X = (X_t)_{t\ge0}$ is given by:

$$X_t = \int_0^t I(s\le r)\,dN_s^{\lambda_0} + \int_0^t I(s>r)\,dN_s^{\lambda_1} \tag{5.2.13}$$

for all $t \ge 0$ where $r \ge 0$. It follows that P_π admits the following decomposition:

$$P_\pi = \pi\,Q_0 + (1-\pi)\int_0^\infty \lambda e^{-\lambda r} Q_r\,dr \tag{5.2.14}$$

which appears to be an elegant tool, for instance, to check that the innovation process $(\widehat{X}_t)_{t\ge0}$ defined above is a martingale under P_π.

Moreover, using (5.2.14) it is straightforwardly verified that the following facts are valid:

(5.2.15) The map $\pi \mapsto V(\pi)$ is concave (continuous) and decreasing on $[0,1]$;

(5.2.16) The stopping time $\tau_* = \inf\{\, t\ge0 \mid \pi_t \ge B_*\,\}$ is optimal in the problem (5.2.5+5.2.7), where B_* is the smallest π from $[0,1]$ satisfying $V(\pi) = 1-\pi$.

Thus $V(\pi) < 1-\pi$ for all $\pi \in [0, B_*\rangle$ and $V(\pi) = 1-\pi$ for all $\pi \in [B_*, 1]$. It should be noted in (5.2.16) that $\pi_t = \varphi_t/(1+\varphi_t)$, and hence by (5.2.9) we see that π_t is a (path-dependent) functional of the process X observed up to time t. Thus, by observing a trajectory of X it is possible to decide when to stop in accordance with the rule τ_* given in (5.2.16).

The question arises, however, to determine the optimal threshold B_* in terms of the four parameters $\lambda_0, \lambda_1, \lambda, c$ as well as to compute the value $V(\pi)$ for $\pi \in [0, B_*\rangle$ (especially for

$\pi = 0$). We tackle these questions by forming a free-boundary problem.

5.3 A free-boundary problem

1. Being aided by the general (optimal stopping) theory of Markov processes (see e.g. [65]), and making use of the preceding facts, we are naturally led to formulate the following *free-boundary problem* for $\pi \mapsto V(\pi)$ and B_* defined above:

(5.3.1) $$(I\!\!LV)(\pi) = -c\pi \quad (0 < \pi < B_*)$$

(5.3.2) $$V(\pi) = 1-\pi \quad (B_* \le \pi \le 1)$$

(5.3.3) $$V(B_*-) = 1-B_* \quad \text{(continuous fit).}$$

In some cases (specified below) the following condition will be satisfied as well:

(5.3.4) $$V'(B_*) = -1 \quad \text{(smooth fit).}$$

However, we will also see below that this condition may fail.

Finally, it is easily verified by passing to the limit for $\pi \downarrow 0$ that each continuous solution $\pi \mapsto V(\pi)$ of the system (5.3.1+5.3.2) must necessarily satisfy:

(5.3.5) $$V'(0+) = 0 \quad \text{(normal entrance)}$$

whenever $V(0+)$ is finite. This condition proves useful in the case when $\lambda_1 < \lambda_0$.

For a similar free-boundary differential-difference problem corresponding to the case $\lambda = 0$ above we refer to [P3].

2. *Solving the free-boundary problem (5.3.1).* It turns out that the case $\lambda_1 < \lambda_0$ is much different from the case $\lambda_1 > \lambda_0$. Thus assume first that $\lambda_1 > \lambda_0$ and consider the equation (5.3.1) on $\langle 0, B]$ for some $0 < B < 1$ given and fixed. Introduce the 'step' function:

(5.3.6) $$S(\pi) = \frac{\lambda_1 \pi}{\lambda_1 \pi + \lambda_0 (1-\pi)}$$

for $\pi \le B$. Observe that $S(\pi) > \pi$ for all $0 < \pi < 1$ and find points $\ldots < B_2 < B_1 < B_0 := B$ such that $S(B_n) = B_{n-1}$ for $n \ge 1$. It is easily verified that:

(5.3.7) $$B_n = \frac{(\lambda_0)^n B}{(\lambda_0)^n B + (\lambda_1)^n (1-B)} \quad (n = 0, 1, \ldots\,) \; .$$

Denote $I_n = \langle B_n, B_{n-1}]$ for $n \ge 1$, and introduce the 'distance' function:

(5.3.8) $$d(\pi, B) = 1 + \left[\log\left(\frac{B}{1-B} \, \frac{1-\pi}{\pi} \right) \Big/ \log\left(\frac{\lambda_1}{\lambda_0} \right) \right]$$

for $\pi \le B$, where $[x]$ denotes the integer part of x . Observe that d is defined to satisfy:

$$\pi \in I_n \iff d(\pi,B) = n \tag{5.3.9}$$

for all $0 < \pi \le B$.

Now consider the equation (5.3.1) first on I_1 upon setting $V(\pi) = 1-\pi$ for $\pi \in \langle B, S(B)]$. This is then a first-order linear differential equation which can be solved explicitly. Imposing a continuity condition at B (which is in agreement with (5.3.3) above) we obtain a *unique* solution $\pi \mapsto V(\pi;B)$ on I_1 . It is possible to verify that the following formula holds:

$$V(\pi;B) = c_1(B)\, V_g(\pi) + V_{p,1}(\pi;B) \quad (\pi \in I_1) \tag{5.3.10}$$

where $\pi \mapsto V_{p,1}(\pi;B)$ is a *particular* solution of the *non-homogeneous* equation in (5.3.1):

$$V_{p,1}(\pi;B) = -\,\frac{\lambda_0(\lambda_1 - c)}{\lambda_1(\lambda_0 + \lambda)}\,\pi + \frac{\lambda_0\lambda_1 + \lambda c}{\lambda_1(\lambda_0+\lambda)} \tag{5.3.11}$$

taken to be bounded, and $\pi \mapsto V_g(\pi)$ is a *general* solution of the *homogeneous* equation in (5.3.1):

$$\begin{aligned} V_g(\pi) &= \frac{(1-\pi)^{\gamma_1}}{|\,\lambda - (\lambda_1 - \lambda_0)\,\pi\,|^{\gamma_0}} \quad , \ \text{if} \ \ \lambda \ne \lambda_1 - \lambda_0 \\ &= (1-\pi)\ \exp\left(\frac{\lambda_1}{(\lambda_1-\lambda_0)(1-\pi)}\right) \quad , \ \text{if} \ \ \lambda = \lambda_1 - \lambda_0 \end{aligned} \tag{5.3.12}$$

where $\gamma_1 = \lambda_1/(\lambda_1 - \lambda_0 - \lambda)$ and $\gamma_0 = (\lambda_0 + \lambda)/(\lambda_1 - \lambda_0 - \lambda)$, and the constant $c_1(B)$ is determined by the continuity condition $V(B-;B) = 1 - B$ leading to:

$$c_1(B) = -\,\frac{1}{V_g(B)}\left(\frac{\lambda_1\lambda + \lambda_0 c}{\lambda_1(\lambda_0+\lambda)}\,B \;-\; \frac{\lambda(\lambda_1 - c)}{\lambda_1(\lambda_0+\lambda)}\right) \tag{5.3.13}$$

where $V_g(B)$ is obtained by replacing π in (5.3.12) by B . [We see from (5.3.11)-(5.3.13) however that the continuity condition at B cannot be met when B equals $\widehat{B}$ from (5.3.16) below unless $\widehat{B}$ equals $\lambda(\lambda_1 - c)/(\lambda\lambda_1 + c\lambda_0)$ from (5.4.5) below (the latter is equivalent to $c = \lambda_1 - \lambda_0 - \lambda$). Thus, if $B = \widehat{B} \ne \lambda(\lambda_1 - c)/(\lambda\lambda_1 + c\lambda_0)$ then there is no solution $\pi \mapsto V(\pi;B)$ on I_1 that satisfies $V(\pi;B) = 1 - \pi$ for $\pi \in \langle B, S(B)]$ and is continuous at B . It turns out, however, that this analytic fact has no significant implication for the solution of (5.2.5+5.2.7).]

Next consider the equation (5.3.1) on I_2 upon using the solution found on I_1 and setting $V(\pi) = c_1(B)\, V_g(\pi) + V_{p,1}(\pi;B)$ for $\pi \in \langle B_1, S(B_1)]$. This is then again a first-order linear differential equation which can be solved explicitly. Imposing a continuity condition over $I_2 \cup I_1$ at B_1 (which is in agreement with (5.2.15) above) we obtain a unique solution $\pi \mapsto V(\pi;B)$ on I_2 . It turns out, however, that the general solution of this equation cannot be expressed in terms of elementary functions (unless $\lambda = 0$ as shown in [P3]) but one needs, for instance, the Gauss hypergeometric function. As these expressions are increasingly complex to record, we omit the explicit formulas in the sequel.

Continuing the preceding procedure by induction as long as possible (considering the equation (5.3.1) on I_n upon using the solution found on I_{n-1} and imposing a continuity condition over $I_n \cup I_{n-1}$ at B_{n-1}) we obtain a unique solution $\pi \mapsto V(\pi;B)$ on I_n given as:

$$V(\pi;B) = c_n(B)\,V_g(\pi) + V_{p,n}(\pi;B) \quad (\pi \in I_n) \tag{5.3.14}$$

where $\pi \mapsto V_{p,n}(\pi;B)$ is a bounded particular solution, $\pi \mapsto V_g(\pi)$ is a general solution given by (5.3.12), and $B \mapsto c_n(B)$ is a function of B (and the four parameters). [We will see however in Theorem 5.4.1 below that in the case $B > \widehat{B} > 0$ with $\widehat{B}$ from (5.3.16) below the solution (5.3.14) exists for $\pi \in \langle \widehat{B}, B]$ but explodes at $\widehat{B}$ unless $B = B_*$.]

The key difference in the case $\lambda_1 < \lambda_0$ is that $S(\pi) < \pi$ for all $0 < \pi < 1$ so that we need to deal with points $B := B_0 < B_1 < B_2 < \ldots$ such that $S(B_n) = B_{n-1}$ for $n \geq 1$. Then the facts (5.3.7)-(5.3.9) remain preserved provided that we set $I_n = [B_{n-1}, B_n\rangle$ for $n \geq 1$. In order to prescribe the initial condition when considering the equation (5.3.1) on I_1 , we can take $B = \varepsilon > 0$ small and make use of (5.3.5) upon setting $V(\pi) = v$ for all $\pi \in [S(B), B\rangle$ where $v \in \langle 0, 1\rangle$ is a given number satisfying $V(B) = v$. Proceeding by induction as earlier (considering the equation (5.3.1) on I_n upon using the solution found on I_{n-1} and imposing a continuity condition over $I_{n-1} \cup I_n$ at B_{n-1}) we obtain a unique solution $\pi \mapsto V(\pi;\varepsilon,v)$ on I_n given as:

$$V(\pi;\varepsilon,v) = c_n(\varepsilon)\,V_g(\pi) + V_{p,n}(\pi;\varepsilon,v) \quad (\pi \in I_n) \tag{5.3.15}$$

where $\pi \mapsto V_{p,n}(\pi;\varepsilon,v)$ is a particular solution, $\pi \mapsto V_g(\pi)$ is a general solution given by (5.3.12), and $\varepsilon \mapsto c_n(\varepsilon)$ is a function of ε (and the four parameters). We shall see in Theorem 5.4.1 below how these solutions can be used to determine the optimal $\pi \mapsto V(\pi)$ and B_* .

3. *Two key facts about the solution.* Both of these facts hold only in the case when $\lambda_1 > \lambda_0$. The first fact to be observed is that:

$$\widehat{B} = \frac{\lambda}{\lambda_1 - \lambda_0} \tag{5.3.16}$$

is a *singularity point* of the equation (5.3.1) whenever $\lambda < \lambda_1 - \lambda_0$. This is clearly seen from (5.3.12) where $V_g(\pi) \to \infty$ for $\pi \to \widehat{B}$. The second fact of interest is that:

$$\widetilde{B} = \frac{\lambda}{\lambda + c} \tag{5.3.17}$$

is a *smooth-fit point* of the system (5.3.1)-(5.3.3) whenever $\lambda_1 > \lambda_0$ and $c \neq \lambda_1 - \lambda_0 - \lambda$, i.e. $V'(\widetilde{B}-;\widetilde{B}) = -1$ in the notation of (5.3.14) above. This can be verified by (5.3.10) using (5.3.11)-(5.3.13). It means that $\widetilde{B}$ is the unique point which in addition to (5.3.1)-(5.3.3) has the power of satisfying the smooth-fit condition (5.3.4).

It may also be noted in the verification above that the equation $V'(B-;B) = -1$ has no solution when $c = \lambda_1 - \lambda_0 - \lambda$ as the only candidate $\bar{B} := \widetilde{B} = \widehat{B}$ satisfies:

$$V'(\bar{B}-;\bar{B}) = -\frac{\lambda_0}{\lambda_1} . \tag{5.3.18}$$

This identity follows readily from (5.3.10)-(5.3.13) upon noticing that $c_1(\bar{B}) = 0$. Thus, when c runs from $+\infty$ to $\lambda_1 - \lambda_0 - \lambda$, the smooth-fit point $\widetilde{B}$ runs from 0 to the singularity point $\widehat{B}$, and once $\widetilde{B}$ has reached $\widehat{B}$ for $c = \lambda_1 - \lambda_0 - \lambda$, the smooth-fit condition (5.3.4) breaks down and gets replaced by the condition (5.3.18) above. We will soon attest below that in all these cases the smooth-fit point $\widetilde{B}$ is actually equal to the optimal-stopping point B_* from (5.2.16) above.

Observe that the equation (5.3.1) has no singularity points when $\lambda_1 < \lambda_0$. This analytic fact reveals a key difference between the two cases.

5.4 Conclusions

In parallel to the two analytic properties displayed above we begin this section by stating the relevant probabilistic properties of the a posteriori probability process.

1. *Sample-path properties of* $(\pi_t)_{t\geq 0}$. First consider the case $\lambda_1 > \lambda_0$. Then from (5.2.11) we see that $(\pi_t)_{t\geq 0}$ can only jump towards 1 (at times of the jumps of the process X). Moreover, the sign of the drift term $\lambda(1-\pi)-(\lambda_1-\lambda_0)\pi(1-\pi) = (\lambda_1-\lambda_0)(\widehat{B}-\pi)(1-\pi)$ is determined by the sign of $\widehat{B}-\pi$. Hence we see that $(\pi_t)_{t\geq 0}$ has a positive drift in $[0,\widehat{B}\rangle$, a negative drift in $\langle \widehat{B},1]$, and a zero drift at $\widehat{B}$. Thus, if $(\pi_t)_{t\geq 0}$ starts or ends up at $\widehat{B}$, it is trapped there until the first jump of the process X occurs. At that time $(\pi_t)_{t\geq 0}$ finally leaves $\widehat{B}$ by jumping towards 1 . This also shows that after once $(\pi_t)_{t\geq 0}$ leaves $[0,\widehat{B}\rangle$ it never comes back. The sample-path behaviour of $(\pi_t)_{t\geq 0}$ when $\lambda_1 > \lambda_0$ is depicted in ***Figure*** 5.1 (Part i) below.

Next consider the case $\lambda_1 < \lambda_0$. Then from (5.2.11) we see that $(\pi_t)_{t\geq 0}$ can only jump towards 0 (at times of the jumps of the process X). Moreover, the sign of the drift term $\lambda(1-\pi)-(\lambda_1-\lambda_0)\pi(1-\pi) = (\lambda+(\lambda_0-\lambda_1)\pi)(1-\pi)$ is always positive. Thus $(\pi_t)_{t\geq 0}$ always moves continuously towards 1 and can only jump towards 0 . The sample-path behaviour of $(\pi_t)_{t\geq 0}$ when $\lambda_1 < \lambda_0$ is depicted in ***Figure*** 5.1 (Part ii) below.

2. *Sample-path behaviour and the principles of smooth and continuous fit.* With a view to (5.2.16), and taking $0 < B < 1$ given and fixed, we shall now examine the manner in which the process $(\pi_t)_{t\geq 0}$ enters $[B,1]$ if starting at $B-d\pi$ where $d\pi$ is infinitesimally small. Our previous analysis then shows the following (see ***Figure*** 5.1 below).

If $\lambda_1 > \lambda_0$ and $B < \widehat{B}$, or $\lambda_1 < \lambda_0$, then $(\pi_t)_{t\geq 0}$ enters $[B,1]$ by passing through B continuously. If, however, $\lambda_1 > \lambda_0$ and $B > \widehat{B}$ then the only way for $(\pi_t)_{t\geq 0}$ to enter $[B,1]$ is by jumping over B . (Jumping exactly at B happens with probability zero.)

The case $\lambda_1 > \lambda_0$ and $B = \widehat{B}$ is special. If starting outside $[B,1]$ then $(\pi_t)_{t\geq 0}$ travels towards $\widehat{B}$ by either moving continuously or by jumping. However, the closer $(\pi_t)_{t\geq 0}$ gets to $\widehat{B}$ the smaller the drift to the right becomes, and if there were no jump over $\widehat{B}$ eventually, the process $(\pi_t)_{t\geq 0}$ would never reach $\widehat{B}$ as the drift to the right tends to zero together with the distance of $(\pi_t)_{t\geq 0}$ to $\widehat{B}$. This fact can be formally verified by analysing the explicit representation of $(\varphi_t)_{t\geq 0}$ in (5.2.9) and using that $\pi_t = \varphi_t/(1+\varphi_t)$ for $t \geq 0$. Thus, in this case as well, the only way for $(\pi_t)_{t\geq 0}$ to enter $[\widehat{B},1]$ after starting at $B-d\pi$ is by jumping over to $\langle \widehat{B},1]$

We will demonstrate below that the sample-path behaviour of the process $(\pi_t)_{t\geq 0}$ during the entrance of $[B_*,1]$ has a precise analytic counterpart in terms of the free-boundary problem (5.3.1). If the process $(\pi_t)_{t\geq 0}$ may enter $[B_*,1]$ by passing through B_* continuously, then the smooth-fit condition (5.3.4) holds at B_* ; if, however, the process $(\pi_t)_{t\geq 0}$ enters $[B_*,1]$ exclusively by jumping over B_* , then the smooth-fit condition (5.3.4) breaks down. In this case the continuous-fit condition (5.3.3) still holds at B_* , and the existence of a singularity point $\widehat{B}$ can be used to determine the optimal B_* as shown below.

3. The preceding considerations may now be summarized as follows.

Theorem 5.4.1

Consider the Poisson disorder problem (5.2.5) and the equivalent optimal-stopping problem (5.2.7) where the process $(\pi_t)_{t \ge 0}$ *from (5.2.6) solves (5.2.11) and* $\lambda_0, \lambda_1, \lambda, c > 0$ *are given and fixed.*

Then there exists $B_* \in \langle 0,1 \rangle$ *such that the stopping time:*

$$\tau_* = \inf \{ \, t \ge 0 \mid \pi_t \ge B_* \, \} \tag{5.4.1}$$

is optimal in (5.2.5) and (5.2.7). Moreover, the optimal cost function $\pi \mapsto V(\pi)$ *from (5.2.5+5.2.7) solves the free-boundary problem (5.3.1)-(5.3.3), and the optimal threshold* B_* *is determined as follows.*

(i): If $\lambda_1 > \lambda_0$ *and* $c > \lambda_1 - \lambda_0 - \lambda$ *, then the smooth-fit condition (5.3.4) holds at* B_* *, and the following explicit formula is valid (cf. [31] and [21]):*

$$B_* \, = \, \frac{\lambda}{\lambda + c} \; . \tag{5.4.2}$$

In this case $B_* < \widehat{B}$ *where* $\widehat{B}$ *is a singularity point of the free-boundary equation (5.3.1) given in (5.3.16) above (see Figure 5.2 below).*

(ii): If $\lambda_1 > \lambda_0$ *and* $c = \lambda_1 - \lambda_0 - \lambda$ *, then the smooth-fit condition breaks down at* B_* *and gets replaced by the condition (5.3.18) above* ($V'(B_*-) = -\lambda_0/\lambda_1$)*. The optimal threshold* B_* *is still given by (5.4.2), and in this case* $B_* = \widehat{B}$ *(see Figure 5.3 below).*

(iii): If $\lambda_1 > \lambda_0$ *and* $c < \lambda_1 - \lambda_0 - \lambda$ *, then the smooth-fit condition does not hold at* B_* *, and the optimal threshold* B_* *is determined as a unique solution in* $\langle \widehat{B}, 1 \rangle$ *of the following equation:*

$$c_{d(\widehat{B}, B_*)}(B_*) \, = \, 0 \tag{5.4.3}$$

where the map $B \mapsto d(\widehat{B}, B)$ *is defined in (5.3.8), and the map* $B \mapsto c_n(B)$ *is defined by (5.3.13) and (5.3.14) above (see Figure 5.4 below). In particular, when* c *satisfies:*

$$\frac{\lambda_1 \lambda_0 \, (\lambda_1 - \lambda_0 - \lambda)}{\lambda_1 \lambda_0 + (\lambda_1 - \lambda_0)(\lambda + \lambda_0)} \le c \, < \, \lambda_1 - \lambda_0 - \lambda \tag{5.4.4}$$

then the following explicit formula is valid:

$$B_* \, = \, \frac{\lambda \, (\lambda_1 - c)}{\lambda \lambda_1 + c \lambda_0} \tag{5.4.5}$$

which in the case $c = \lambda_1 - \lambda_0 - \lambda$ *reduces again to (5.4.2) above.*

In the cases (i)-(iii) the optimal cost function $\pi \mapsto V(\pi)$ *from (5.2.5+5.2.7) is given by (5.3.14) with* B_* *in place of* B *for all* $0 < \pi \le B_*$ *(with* $V(0) = V(0+)$ *) and* $V(\pi) = 1 - \pi$ *for* $B_* \le \pi \le 1$ *.*

(iv): If $\lambda_1 < \lambda_0$ *then the smooth-fit condition holds at* B_* *, and the optimal threshold* B_* *can be determined using the normal entrance condition (5.3.5) as follows (see Figure 5.5 below). For* $\varepsilon > 0$ *small let* v_ε *denote a unique number in* $\langle 0,1 \rangle$ *for which the map* $\pi \mapsto V(\pi; \varepsilon, v_\varepsilon)$ *from (5.3.15) hits the map* $\pi \mapsto 1 - \pi$ *smoothly at some* B_*^ε *from* $\langle 0,1 \rangle$ *. Then we have:*

$$(5.4.6) \qquad B_* = \lim_{\varepsilon \downarrow 0} B_*^{\varepsilon}$$

$$(5.4.7) \qquad V(\pi) = \lim_{\varepsilon \downarrow 0} V(\pi; \varepsilon, v_\varepsilon)$$

for all $0 < \pi \le B_*$ (*with* $V(0) = V(0+)$) *and* $V(\pi) = 1 - \pi$ *for* $B_* \le \pi \le 1$.

Proof. We have already established in (5.2.16) above that τ_* from (5.4.1) is optimal in (5.2.5) and (5.2.7) for some $B_* \in [0,1]$ to be found. It thus follows by the strong Markov property of the process $(\pi_t)_{t \ge 0}$ together with (5.2.15) above that the optimal cost function $\pi \mapsto V(\pi)$ from (5.2.5+5.2.7) solves the free-boundary problem (5.3.1)-(5.3.3). Some of these facts will also be reproved below.

First consider the case $\lambda_1 > \lambda_0$. In Subsection 5.3.2 above it was shown that for each given and fixed $B \in \langle 0, \widehat{B} \rangle$ the problem (5.3.1)-(5.3.3) with B in place of B_* has a unique continuous solution given by the formula (5.3.14). Moreover, this solution is (at least) C^1 everywhere but possibly at B where it is (at least) C^0. As explained following (5.3.13) above, these facts also hold for $B = \widehat{B}$ when $\widehat{B}$ equals $\lambda(\lambda_1 - c)/(\lambda\lambda_1 + c\lambda_0)$ from (5.4.5) above. We will now show how the optimal threshold B_* is determined among all these candidates B when $c \ge \lambda_1 - \lambda_0 - \lambda$.

(i)+(ii): Since the innovation process $\widehat{X}_t = X_t - \int_0^t (\lambda_1 \pi_{s-} + \lambda_0(1 - \pi_{s-}))\, ds$ is a martingale under P_π with respect to $(\mathcal{F}_t^X)_{t \ge 0}$, it follows by (5.2.11) that:

$$(5.4.8) \qquad \pi_t = \pi + \lambda \int_0^t (1 - \pi_{s-})\, ds + M_t$$

where $M = (M_t)_{t \ge 0}$ is a martingale under P_π with respect to $(\mathcal{F}_t^X)_{t \ge 0}$. Hence by the optional sampling theorem we easily find:

$$(5.4.9) \qquad E_\pi \left((1 - \pi_\tau) + c \int_0^\tau \pi_t\, dt \right) = (1 - \pi) + (\lambda + c)\, E_\pi \left(\int_0^\tau \left(\pi_t - \frac{\lambda}{\lambda + c} \right) dt \right)$$

for all stopping times τ of $(\pi_t)_{t \ge 0}$. Recalling the sample-path behaviour of $(\pi_t)_{t \ge 0}$ in the case $\lambda_1 > \lambda_0$ as displayed in Subsection 5.4.1 above (cf. *Figure* 5.1 (Part i) below), and the definition of $V(\pi)$ in (5.2.7) together with the fact that $\widetilde{B} = \lambda/(\lambda + c) \le \widehat{B}$ when $c \ge \lambda_1 - \lambda_0 - \lambda$, we clearly see from (5.4.9) that it is never optimal to stop $(\pi_t)_{t \ge 0}$ in $[0, \widetilde{B} \rangle$, as well as that $(\pi_t)_{t \ge 0}$ must be stopped immediately after entering $[\widetilde{B}, 1]$ as it will never return to the 'favourable' region $[0, \widetilde{B} \rangle$ again. This proves that $\widetilde{B}$ equals the optimal threshold B_* , i.e. that τ_* from (5.4.1) with B_* from (5.4.2) is optimal in (5.2.5) and (5.2.7). The claim about the breakdown of the smooth-fit condition (5.3.4) when $c = \lambda_1 - \lambda_0 - \lambda$ has been already established in the paragraph containing (5.3.18) above (cf. *Figure* 5.3 below).

(iii): It was shown in Subsection 5.3.2 above that for each given and fixed $B \in \langle \widehat{B}, 1 \rangle$ the problem (5.3.1)-(5.3.3) with B in place of B_* has a unique continuous solution on $\langle \widehat{B}, 1]$ given by the formula (5.3.14). We will now show that there exists a unique point $B_* \in \langle \widehat{B}, 1 \rangle$ such that $\lim_{\pi \downarrow \widehat{B}} V(\pi; B) = \pm\infty$ if $B \in \langle \widehat{B}, B_* \rangle \cup \langle B_*, 1 \rangle$ and $\lim_{\pi \downarrow \widehat{B}} V(\pi; B_*)$ is finite. This point is the optimal threshold, i.e. the stopping time τ_* from (5.4.1) is optimal in (5.2.5) and (5.2.7). Moreover, the point B_* can be characterized as a unique solution of the equation (5.4.3) in $\langle \widehat{B}, 1 \rangle$.

In order to verify the preceding claims we will first state the following observation which proves useful. Setting $g(\pi) = 1 - \pi$ for $0 < \pi < 1$ we have:

$$(\mathbb{L}g)(\pi) \geq -c\pi \iff \pi \geq \widetilde{B} \tag{5.4.10}$$

where $\widetilde{B}$ is given in (5.3.17). This is verified straightforwardly using (5.2.12).

Now since $\widehat{B}$ is a singularity point of the equation (5.3.1) (recall our discussion in Subsection 5.3.3 above), and moreover $\pi \mapsto V(\pi)$ from (5.2.5+5.2.7) solves (5.3.1)-(5.3.3), we see that the optimal threshold B_* from (5.2.16) must satisfy (5.4.3). This is due to the fact that a particular solution $\pi \mapsto V_{p,n}(\pi; B_*)$ for $n = d(\widehat{B}, B_*)$ in (5.3.14) above is taken bounded. The key remaining fact to be established is that there cannot be two (or more) points in $\langle \widehat{B}, 1 \rangle$ satisfying (5.4.3).

Assume on the contrary that there are two such points B_1 and B_2. We may however assume that both B_1 and B_2 are larger than $\widetilde{B}$ since for $B \in \langle \widehat{B}, \widetilde{B} \rangle$ the solution $\pi \mapsto V(\pi; B)$ is ruled out by the fact that $V(\pi; B) > 1 - \pi$ for $\pi \in \langle B - \varepsilon, B \rangle$ with $\varepsilon > 0$ small. This fact is verified directly using (5.3.10)-(5.3.13). Thus, each map $\pi \mapsto V(\pi; B_i)$ solves (5.3.1)-(5.3.3) on $\langle 0, B_i]$ and is continuous (bounded) at $\widehat{B}$ for $i = 1, 2$. Since $S(\pi) > \pi$ for all $0 < \pi < 1$ when $\lambda_1 > \lambda_0$, it follows easily from (5.2.12) that each solution $\pi \mapsto V(\pi; B_i)$ of (5.3.1)-(5.3.3) must also satisfy $-\infty < V(0+; B_i) < +\infty$ for $i = 1, 2$.

In order to make use of the preceding fact we shall set $h_\beta(\pi) = (1 + (\beta - 1)\widehat{B}) - \beta\pi$ for $0 \leq \pi \leq \widehat{B}$ and $h_\beta(\pi) = 1 - \pi$ for $\widehat{B} \leq \pi \leq 1$. Since both maps $\pi \mapsto V(\pi; B_i)$ are bounded on $\langle 0, \widehat{B} \rangle$ we can fix $\beta > 0$ large enough so that $V(\pi; B_i) \leq h_\beta(\pi)$ for all $0 < \pi \leq \widehat{B}$ and $i = 1, 2$. Consider then the auxiliary optimal stopping problem:

$$W(\pi) := \inf_{\tau} E_\pi \Big(h_\beta(\pi_\tau) + c \int_0^\tau \pi_t \, dt \Big) \tag{5.4.11}$$

where the supremum is taken over all stopping times τ of $(\pi_t)_{t \geq 0}$. Extend the map $\pi \mapsto V(\pi; B_i)$ on $[B_i, 1]$ by setting $V(\pi; B_i) = 1 - \pi$ for $B_i \leq \pi \leq 1$ and denote the resulting (continuous) map on $[0, 1]$ by $\pi \mapsto V_i(\pi)$ for $i = 1, 2$. Then $\pi \mapsto V_i(\pi)$ satisfies (5.3.1)-(5.3.3), and since $B_i \geq \widetilde{B}$, we see by means of (5.4.10) that the following condition is also satisfied:

$$(\mathbb{L}V_i)(\pi) \geq -c\pi \tag{5.4.12}$$

for $\pi \in [B_i, 1]$ and $i = 1, 2$. We will now show that the preceding two facts have the power of implying that $V_i(\pi) = W(\pi)$ for all $\pi \in [0, 1]$ with either $i \in \{1, 2\}$ given and fixed.

It follows by Itô formula that:

$$V_i(\pi_t) = V_i(\pi) + \int_0^t (\mathbb{L}V_i)(\pi_{s-}) \, ds + M_t \tag{5.4.13}$$

where $M = (M_t)_{t \geq 0}$ is a martingale (under P_π) given by:

$$M_t = \int_0^t \Big(V_i\big(\pi_{s-} + \Delta\pi_s\big) - V_i(\pi_{s-}) \Big) \, d\widehat{X}_s \tag{5.4.14}$$

and $\widehat{X}_t = X_t - \int_0^t (\lambda_1 \pi_{s-} + \lambda_0 (1 - \pi_{s-})) \, ds$ is the innovation process. By the optional sampling theorem it follows from (5.4.13) using (5.4.12) and the fact that $V_i(\pi) \le h_\beta(\pi)$ for all $\pi \in [0,1]$ that $V_i(\pi) \le W(\pi)$ for all $\pi \in [0,1]$. Moreover, defining $\tau_i = \inf\{\, t \ge 0 \mid \pi_t \ge B_i \,\}$ it is easily seen by (5.4.8) for instance that $E_\pi(\tau_i) < \infty$. Using then that $\pi \mapsto V_i(\pi)$ is bounded on $[0,1]$, it follows easily by the optional sampling theorem that $E_\pi(M_{\tau_i}) = 0$. Since moreover $V_i(\pi_{\tau_i}) = h_\beta(\pi_{\tau_i})$ and $(I\!\!LV_i)(\pi_{s-}) = -c\pi_{s-}$ for all $s \le \tau_i$, we see from (5.4.13) that the inequality $V_i(\pi) \le W(\pi)$ derived above is actually equality for all $\pi \in [0,1]$. This proves that $V(\pi; B_1) = V(\pi; B_2)$ for all $\pi \in [0,1]$, or in other words, that there cannot be more than one point B_* in $\langle \widehat{B}, 1 \rangle$ satisfying (5.4.3). Thus, there is only one solution $\pi \mapsto V(\pi)$ of (5.3.1)-(5.3.3) which is finite at $\widehat{B}$ (see *Figure* 5.4 below), and the proof of the claim is complete.

(iv): It was shown in Subsection 5.3.2 above that the map $\pi \mapsto V(\pi;\varepsilon,v)$ from (5.3.15) is a unique continuous solution of the equation $(I\!\!LV)(\pi) = -c\pi$ for $\varepsilon < \pi < 1$ satisfying $V(\pi) = v$ for all $\pi \in [S(\varepsilon), \varepsilon]$. It can be checked using (5.3.12) that:

$$V_{p,1}(\pi;\varepsilon,v) = \frac{c\lambda_0}{\lambda_1(\lambda_0+\lambda)}\,\pi + \frac{c\lambda}{\lambda_1(\lambda_0+\lambda)} + v \tag{5.4.15}$$

$$c_1(\varepsilon) = -\frac{1}{V_g(\varepsilon)}\left(\frac{c\lambda_0}{\lambda_1(\lambda_0+\lambda)}\,\varepsilon + \frac{c\lambda}{\lambda_1(\lambda_0+\lambda)} \right) \tag{5.4.16}$$

for $\pi \in I_1 = [\varepsilon, \varepsilon_1\rangle$ where $S(\varepsilon_1) = \varepsilon$. Moreover, it may be noted directly from (5.2.12) above that $I\!\!L(f+c) = I\!\!L(f)$ for every constant c , and thus $V(\pi;\varepsilon,v) = V(\pi;\varepsilon,0) + v$ for all $\pi \in [S(\varepsilon), 1\rangle$. Consequently, the two maps $\pi \mapsto V(\pi;\varepsilon,v')$ and $\pi \mapsto V(\pi;\varepsilon,v'')$ do not intersect in $[S(\varepsilon), 1\rangle$ when v' and v'' are different.

Each map $\pi \mapsto V(\pi;\varepsilon,v)$ is concave on $[S(\varepsilon), 1\rangle$. This fact can be proved by a probabilistic argument using (5.2.14) upon considering the auxiliary optimal stopping problem (5.4.11) where the map $\pi \mapsto h_\beta(\pi)$ is replaced by the concave map $h_v(\pi) = v \wedge (1-\pi)$. [It is a matter of fact that $\pi \mapsto W(\pi)$ from (5.4.11) is concave on $[0,1]$ whenever $\pi \mapsto h_\beta(\pi)$ is so.] Moreover, using (5.3.12)+(5.4.15)+(5.4.16) in (5.3.15) with $n=1$ it is possible to see that for v close to 0 we have $V(\pi;\varepsilon,v) < 0$ for some $\pi > \varepsilon$, and for v close to 1 we have $V(\pi;\varepsilon,v) > 1-\pi$ for some $\pi > \varepsilon$ (see *Figure* 5.5 below). Thus a simple concavity argument implies the existence of a unique point $B_*^\varepsilon \in \langle 0,1 \rangle$ at which $\pi \mapsto V(\pi;\varepsilon,v_\varepsilon)$ for some $v_\varepsilon \in \langle 0,1 \rangle$ hits $\pi \mapsto 1-\pi$ smoothly. The key non-trivial point in the verification that $V(\pi;\varepsilon,v_\varepsilon)$ equals the value function $W(\pi)$ of the optimal stopping problem (5.4.11) with $\pi \mapsto h_{v_\varepsilon}(\pi)$ in place of $\pi \mapsto h_\beta(\pi)$ is to establish that $(I\!\!L(V(\,\cdot\,;\varepsilon,v_\varepsilon)))(\pi) \ge -c\pi$ for all $\pi \in \langle B_*^\varepsilon, S^{-1}(B_*^\varepsilon) \rangle$. Since B_*^ε is a smooth-fit point, however, this can be done using the same method which we applied in part 3 of the proof of Theorem 2.1 in [P3]. Moreover, when $\varepsilon \downarrow 0$ then clearly (5.4.6) and (5.4.7) are valid (recall (5.2.15) and (5.3.5) above), and the proof of the theorem is complete. □

Concluding the chapter we would like to mention that the fixed false-alarm formulation of the Poisson disorder problem (cf. [65; p.205]) raises some new interesting questions not present in the Wiener process version of the same problem.

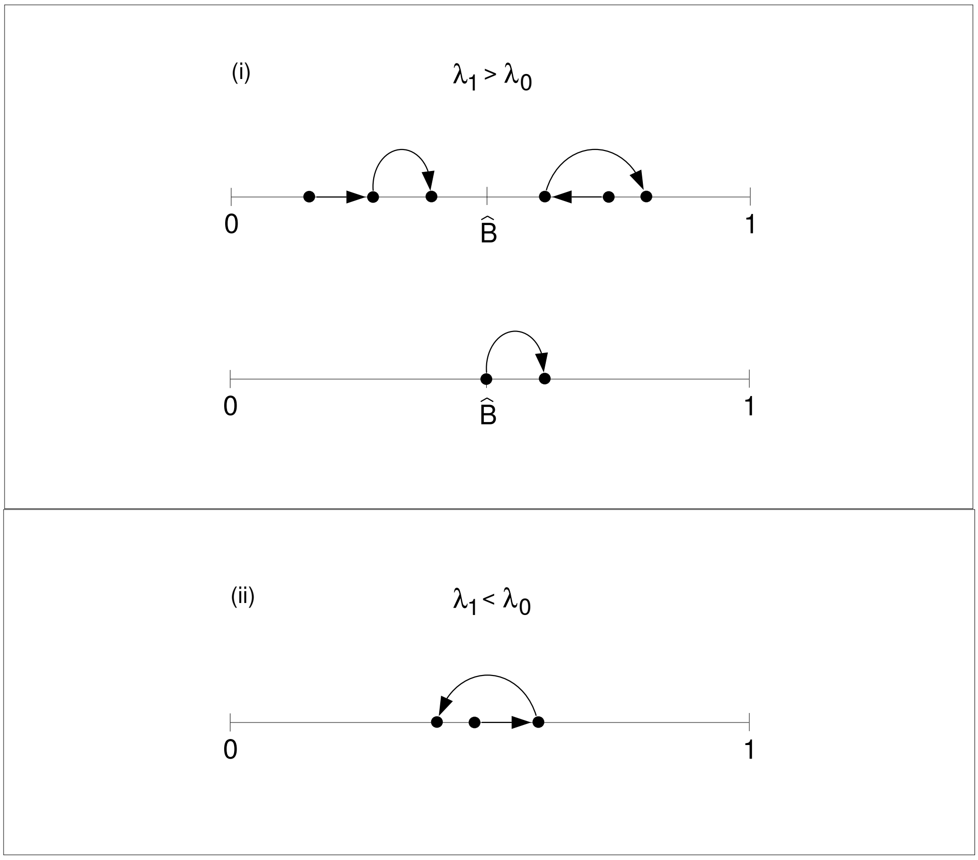


Figure 5.1. Sample-path properties of the a posteriori probability process $(\pi_t)_{t \geq 0}$ from (5.2.6+5.2.11). The point $\widehat{B}$ is a singularity point (5.3.16) of the free-boundary equation (5.3.1).

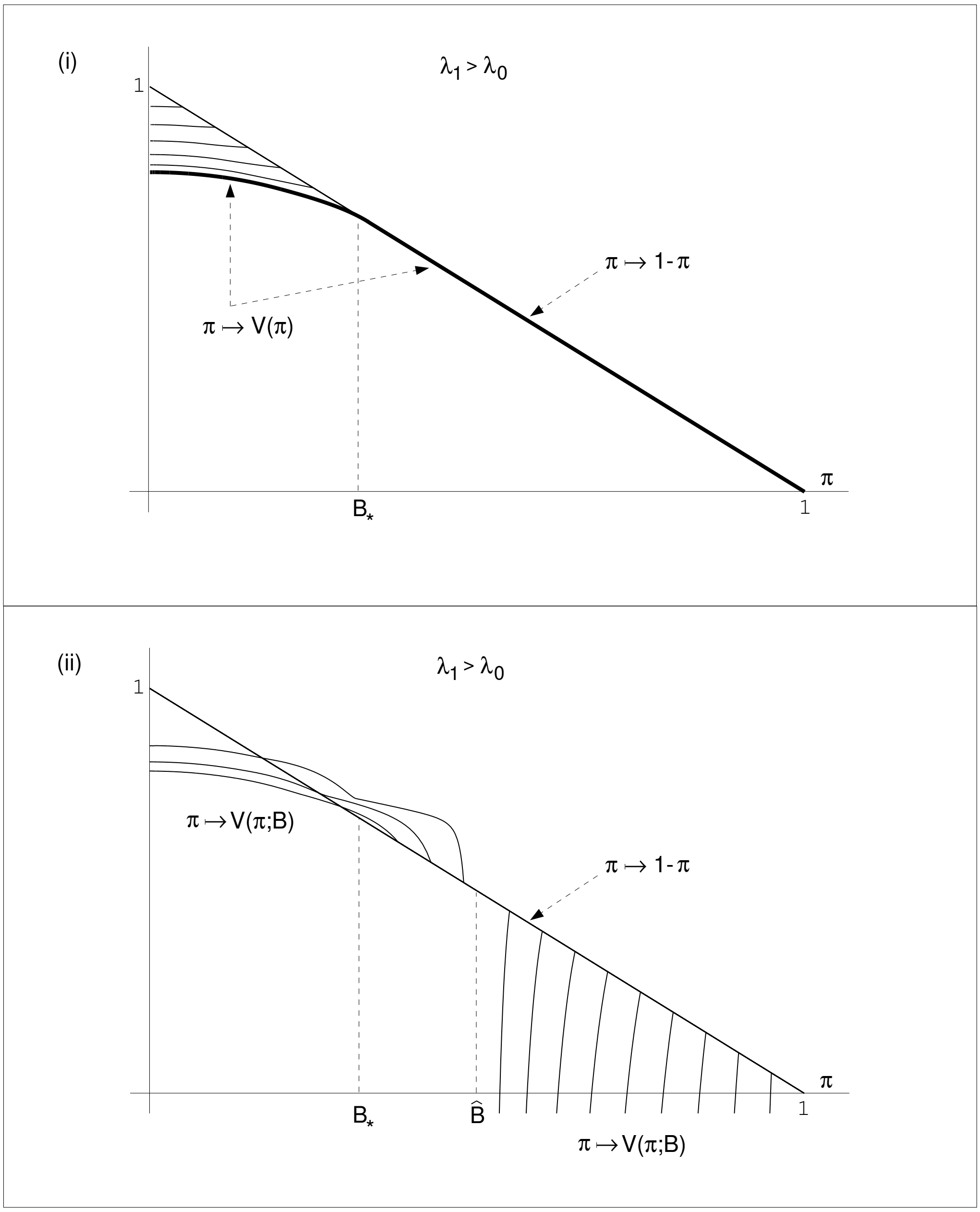


Figure 5.2. A computer drawing of the maps $\pi \mapsto V(\pi;B)$ from (5.3.14) for different B from $\langle 0,1 \rangle$ in the case $\lambda_1 = 4$, $\lambda_0 = 2$, $\lambda = 1$, $c = 2$. The singularity point $\widehat{B}$ from (5.3.16) equals $1/2$, and the smooth-fit point $\widetilde{B}$ from (5.3.17) equals $1/3$. The optimal threshold B_* coincides with the smooth-fit point $\widetilde{B}$. The optimal cost function $\pi \mapsto V(\pi)$ from (5.2.5+5.2.7) equals $\pi \mapsto V(\pi;B_*)$ for $0 \le \pi \le B_*$ and $1-\pi$ for $B_* \le \pi \le 1$. (This is presented in part (i) above.) The solutions $\pi \mapsto V(\pi;B)$ for $B > B_*$ are ruled out since they fail to satisfy $0 \le V(\pi) \le 1-\pi$ for all $\pi \in [0,1]$. (This is shown in part (ii) above.) The general case $\lambda_1 > \lambda_0$ with $c > \lambda_1 - \lambda_0 - \lambda$ looks very much the same.

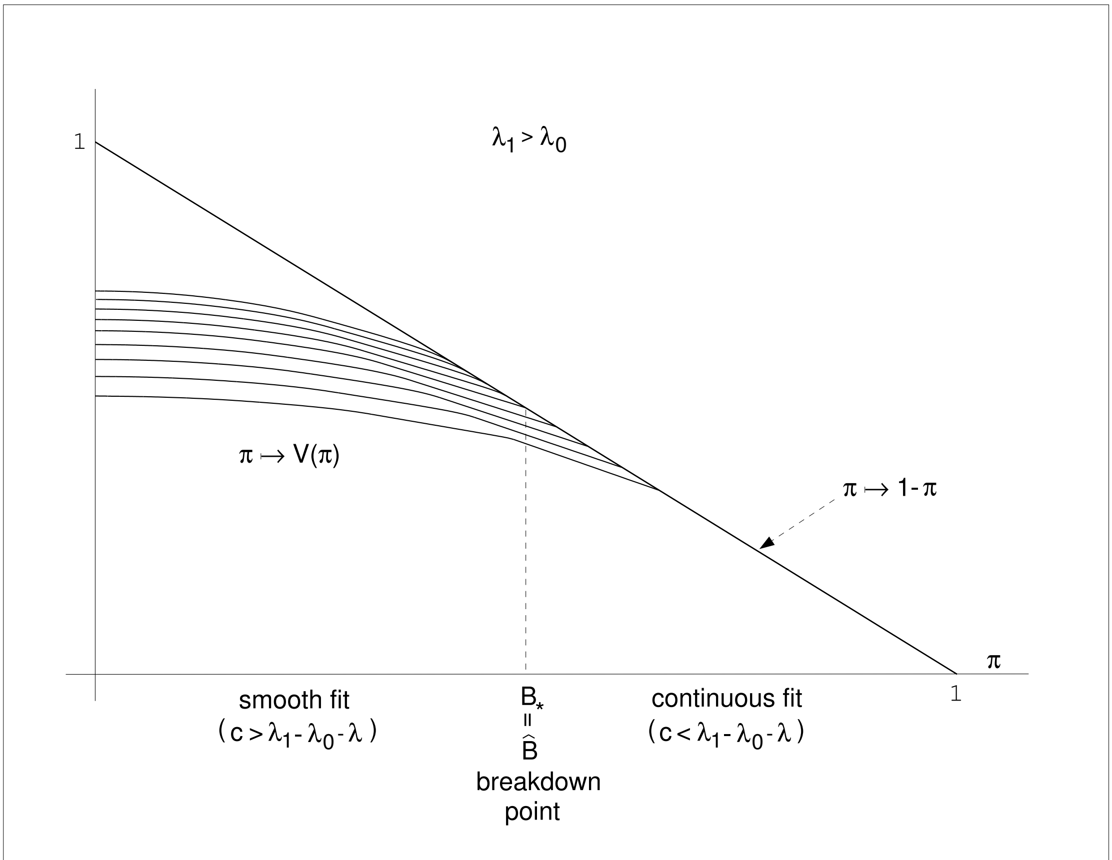
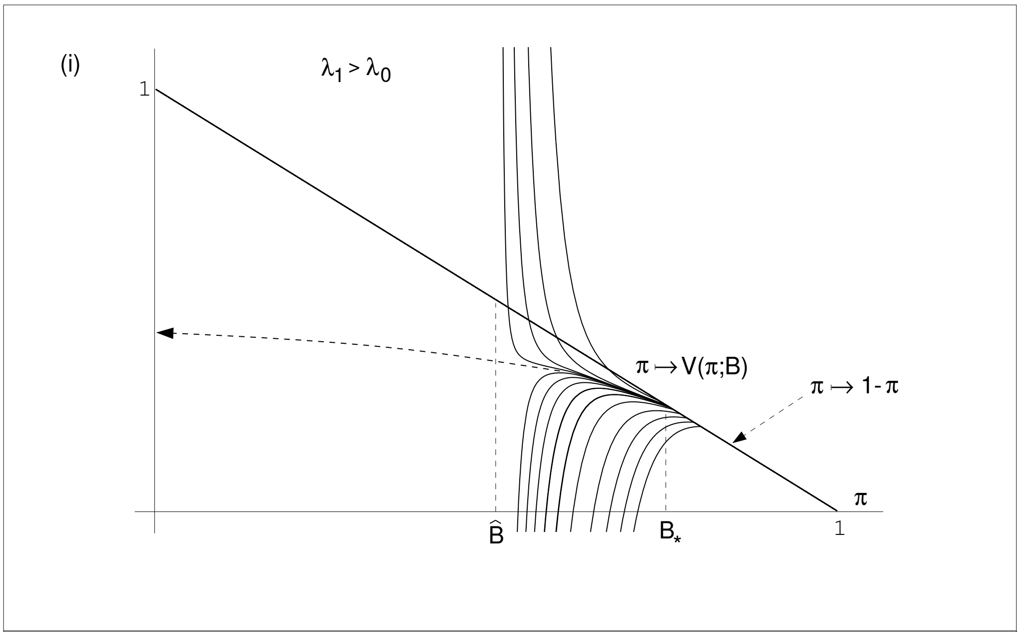


Figure 5.3. A computer drawing of the optimal cost functions $\pi \mapsto V(\pi)$ from (5.2.5+5.2.7) in the case $\lambda_1 = 4$, $\lambda_0 = 2$, $\lambda = 1$ and $c = 1.4,\ 1.3,\ 1.2,\ 1.1,\ 1,\ 0.9,\ 0.8,\ 0.7,\ 0.6$. The given $V(\pi)$ equals $V(\pi; B_*)$ from (5.3.14) for all $0 < \pi \le B_*$ where B_* as a function of c is given by (5.4.2) and (5.4.5). The smooth-fit condition (5.3.4) holds in the cases $c = 1.4,\ 1.3,\ 1.2,\ 1.1$. The point $c = 1$ is a breakdown point when the optimal threshold B_* equals the singularity point $\widehat{B}$ from (5.3.16), and the smooth-fit condition gets replaced by the condition (5.3.18) with $\bar{B} = B_* = \widehat{B} = 0.5$ in this case. For $c = 0.9,\ 0.8,\ 0.7,\ 0.6$ the smooth-fit condition (5.3.4) does not hold. In these cases the continuous-fit condition (5.3.3) is satisfied. Moreover, numerical computations suggest that the mapping $B_* \mapsto V'(B_*-; B_*)$ which equals -1 for $0 < B_* < \widehat{B}$ and jumps to $-\lambda_0/\lambda_1 = -0.5$ for $B_* = \widehat{B}$ is decreasing on $[\widehat{B}, 1\rangle$ and tends to a value slightly larger than -0.6 when $B_* \uparrow 1$ that is $c \downarrow 0$. The general case $\lambda_1 > \lambda_0$ looks very much the same.



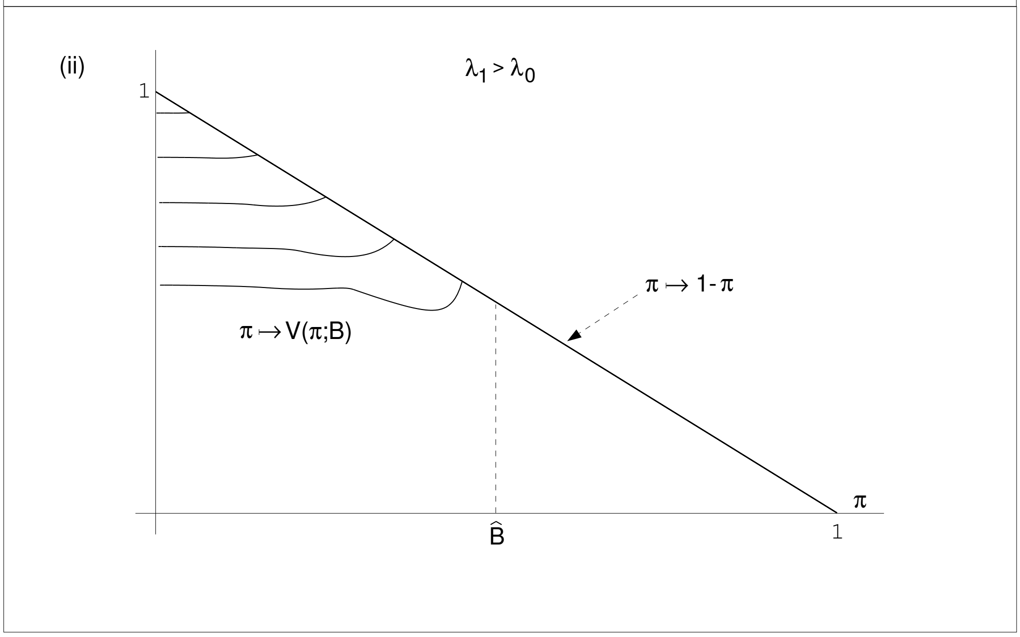


Figure 5.4. A computer drawing of the maps $\pi \mapsto V(\pi;B)$ from (5.3.14) for different B from $\langle 0,1\rangle$ in the case $\lambda_1 = 4$, $\lambda_0 = 2$, $\lambda = 1$, $c = 2/5$. The singularity point $\widehat{B}$ from (5.3.16) equals $1/2$. The optimal threshold B_* can be determined from the fact that all solutions $\pi \mapsto V(\pi;B)$ for $B > B_*$ hit zero for some $\pi > \widehat{B}$, and all solutions $\pi \mapsto V(\pi;B)$ for $B < B_*$ hit $1-\pi$ for some $\pi > \widehat{B}$. (This is shown in part (i) above.) A simple numerical method based on the preceding fact suggests the following estimates $0.750 < B_* < 0.752$. The optimal cost function $\pi \mapsto V(\pi)$ from (5.2.5+5.2.7) equals $\pi \mapsto V(\pi;B_*)$ for $0 \le \pi \le B_*$ and $1-\pi$ for $B_* \le \pi \le 1$. The solutions $\pi \mapsto V(\pi;B)$ for $B \le \widehat{B}$ are ruled out since they fail to be concave. (This is shown in part (ii) above.) The general case $\lambda_1 > \lambda_0$ with $c < \lambda_1 - \lambda_0 - \lambda$ looks very much the same.

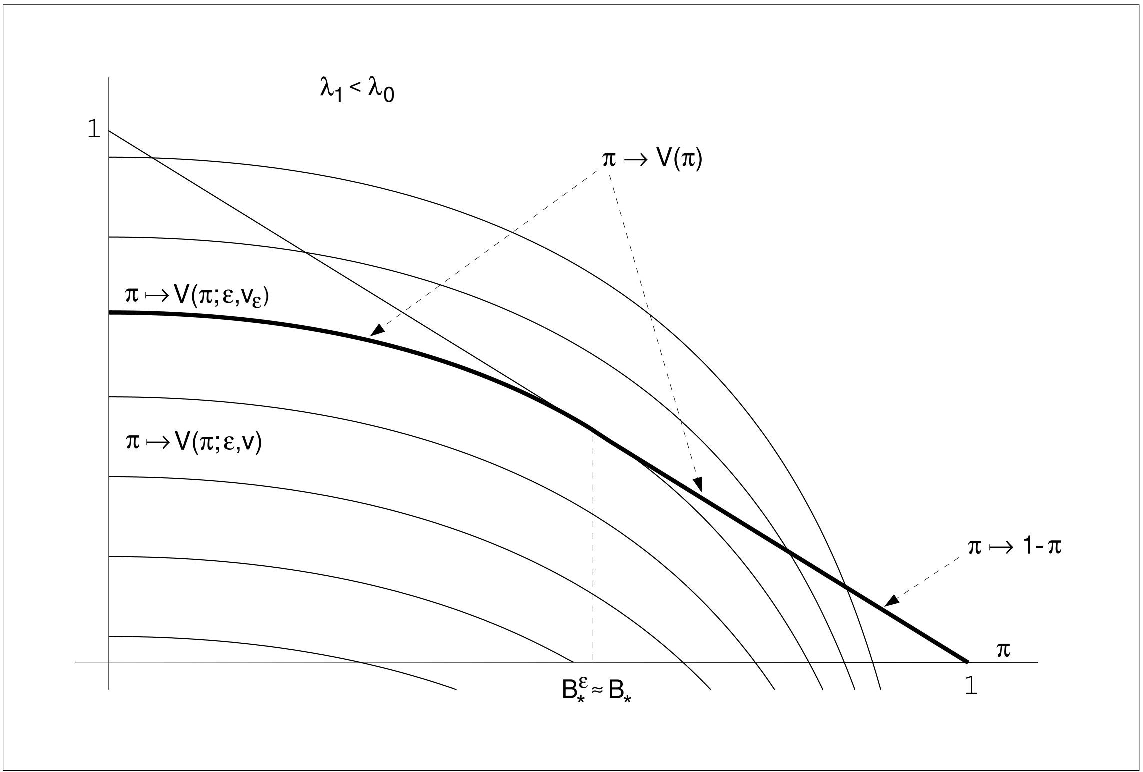


Figure 5.5. A computer drawing of the maps $\pi \mapsto V(\pi;\varepsilon,v)$ from (5.3.15) for different v from $\langle 0,1\rangle$ with $\varepsilon = 0.001$ in the case $\lambda_1 = 2$, $\lambda_0 = 4$, $\lambda = 1$, $c = 1$. For each $\varepsilon > 0$ there is a unique number $v_\varepsilon \in \langle 0,1\rangle$ such that the map $\pi \mapsto V(\pi;\varepsilon,v_\varepsilon)$ hits the map $\pi \mapsto 1-\pi$ smoothly at some $B_*^\varepsilon \in \langle 0,1\rangle$. Letting $\varepsilon \downarrow 0$ we obtain $B_*^\varepsilon \to B_*$ and $V(\pi;\varepsilon,v_\varepsilon) \to V(\pi)$ for all $\pi \in [0,1]$ where B_* is the optimal threshold from (5.2.16) and $\pi \mapsto V(\pi)$ is the optimal cost function from (5.2.5+5.2.7).

REFERENCES

[1] ARROW, K. J. BLACKWELL, D. *and* GIRSHICK, M. A. (1949). Bayes and minimax solutions of sequential decision problems. *Econometrica* 17 (213-244).

[2] AZEMA, J. *and* YOR, M. (1979). Une solution simple au probleme de Skorokhod. *Sem. Probab.* XIII, *Lecture Notes in Math.* 721, Springer-Verlag (90-115).

[3] AZEMA, J. *and* YOR, M. (1979). Le probleme de Skorokhod: Complements a l'expose precedent. *Sem. Probab.* XIII, *Lecture Notes in Math.* 721, Springer-Verlag (625-633).

[4] BATHER, J. *and* CHERNOFF, H. (1967). Sequential decisions in the control of a spaceship. *Proc. 5th Berkeley Sympos. Math. Statist. Probab.* Vol III, Univ. California Press (181-207).

[5] BATHER, J. *and* CHERNOFF, H. (1967). Sequential decisions in the control of a spaceship (finite fuel). *J. Appl. Probab.* 4 (584-604).

[6] BELLMAN, R. (1952). On the theory of dynamic programming. *Proc. Nat. Acad. Sci. U.S.A.* 38 (716-719).

[7] BELLMAN, R. (1957). *Dynamic Programming*. Princeton Univ. Press.

[8] BENEŠ, V. E. SHEPP, L. A. *and* WITSENHAUSEN, H. S. (1980). Some solvable stochastic control problems. *Stochastics* 4 (39-83).

[9] BHAT, B. R. (1988). Optimal properties of SPRT for some stochastic processes. *Contemp. Math.* 80 (285-299).

[10] BLACKWELL, D. *and* GIRSHICK, M. A. (1954). *Theory of Games and Statistical Decisions*. John Wiley and Sons, New York.

[11] BLISS, G. A. (1925). *Calculus of Variations*. Carus Math. Monographs No. 1, Math. Assoc. Amer.

[12] BLISS, G. A. (1946). *Lectures on the Calculus of Variations*. Univ. Chicago Press.

[13] BOLTYANSKI, V. G. GAMKRELIDZE, R. V. *and* PONTRYAGIN, L. S. (1956). On the theory of optimal processes. *Dokl. Akad. Nauk SSSR (N.S.)* 110 (7-10).

[14] BOYCE, W. M. (1970). Stopping rules for selling bonds. *Bell J. Econom. Manage. Sci.* 1 (27-53).

[15] BURKHOLDER, D. L. (1991). Explorations in martingale theory and its applications. *Ecole d'Eté de Probabilités de Saint-Flour* XIX-1989, *Lecture Notes in Math.* 1464, Springer-Verlag (1-66).

[16] CARATHÉODORI, C. (1965). *Calculus of Variations and Partial Differential Equations of the First Order. Part I: Partial Differential Equations of the First Order*. Holden-Day.

[17] CARATHÉODORI, C. (1967). *Calculus of Variations and Partial Differential Equations of the First Order. Part II: Calculus of Variations*. Holden-Day.

[18] CHERNOFF, H. (1961). Sequential tests for the mean of a normal distribution. *Proc. 4th Berkeley Sympos. Math. Statist. Probab.* Vol I, Univ. California Press (79-91).

[19] CHOW, Y. S. ROBBINS, H. *and* SIEGMUND, D. (1971). *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin, Boston, Mass.

[20] COX, D. C. (1984). Some sharp martingale inequalities related to Doob's inequality. *IMS Lecture Notes Monogr. Ser.* 5 (78-83).

[21] DAVIS, M. H. A. (1976). A note on the Poisson disorder problem. *Banach Center Publ.* 1 (65-72).

[22] DUBINS, L. E. *and* GILAT, D. (1978). On the distribution of maxima of martingales. *Proc. Amer. Math. Soc.* 68 (337-338).

[23] DUBINS, L. E. *and* SCHWARZ, G. (1988). A sharp inequality for sub-martingales and stopping times. *Astérisque* 157-158 (129-145).

[24] DUBINS, L. E. SHEPP, L. A. *and* SHIRYAEV, A. N. (1993). Optimal stopping rules and maximal inequalities for Bessel processes. *Theory Probab. Appl.* 38 (226-261).

[25] DVORETZKY, A. KIEFER, J. *and* WOLFOWITZ, J. (1953). Sequential decision problems for processes with continuous time parameter. Testing hypotheses. *Ann. Math. Statist.* 24 (254-264).

[26] DYNKIN, E. B. (1963). The optimum choice of the instant for stopping a Markov process. *Soviet Math. Dokl.* 4 (627-629).

[27] FLEMING, W. H. *and* RISHEL, R. W. (1975). *Deterministic and Stochastic Optimal Control*. Springer-Verlag.

[28] FLEMING, W. H. *and* SONER, H. M. (1993). *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag.

[29] FRIEDMAN, A. (1964). *Partial Differential Equations of Parabolic Type*. Prentice-Hall.

[30] FRIEDMAN, A. (2000). Free boundary problems in science and technology. *Notices Amer. Math. Soc.* 47 (854-861).

[31] GAL'CHUK, L. I. *and* ROZOVSKII, B. L. (1971). The "disorder" problem for a Poisson process. *Theory Probab. Appl.* 16 (712-716).

[32] GILAT, D. (1986). The best bound in the $L \log L$ inequality of Hardy and Littlewood and its martingale counterpart. *Proc. Amer. Math. Soc.* 97 (429-436).

[33] GILAT, D. (1988). On the ratio of the expected maximum of a martingale and the L_p-norm of its last term. *Israel J. Math.* 63 (270-280).

[34] GILBERT, J. P. *and* MOSTELLER, F. (1966). Recognizing the maximum of a sequence. *J. Amer. Statist. Assoc.* 61 (35-73).

[35] GOLDSTEIN, H. (1951). *Classical Mechanics*. Addison-Wesley Press.

[36] GRIFFEATH, D. *and* SNELL, J. L. (1974). Optimal stopping in the stock market. *Ann. Probab.* 2 (1-13).

[37] GRIGELIONIS, B. I. *and* SHIRYAEV, A. N. (1966). On Stefan's problem and optimal stopping rules for Markov processes. *Theory Probab. Appl.* 11 (541-558).

[38] IRLE, A. *and* SCHMITZ, N. (1984). On the optimality of the SPRT for processes with continuous parameter. *Math. Operationsforsch. Statist. Ser. Statist.* 15 (91-104).

[39] JACKA, S. D. (1988). Doob's inequalities revisited: A maximal H^1-embedding. *Stochastic Process. Appl.* 29 (281-290).

[40] JACKA, S. D. (1991). Optimal stopping and best constants for Doob-like inequalities I: The case $p = 1$. *Ann. Probab.* 19 (1798-1821).

[41] JACOD, J. *and* SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes*. Springer-Verlag.

[42] KAKUTANI, S. (1944). Two-dimensional Brownian motion and harmonic functions. *Proc. Imp. Acad. Tokyo* 20 (706-714).

[43] KARATZAS, I. *and* SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*. Springer-Verlag.

[44] KARATZAS, I. *and* SHREVE, S. (1998). *Methods of Mathematical Finance*. Springer-Verlag.

[45] KARLIN, S. (1962). Stochastic models and optimal policy for selling an asset. *Studies Appl. Probab. Manage. Sci.*, Stanford Univ. Press (148-158).

[46] KARLIN, S. *and* TAYLOR, H. M. (1981). *A Second Course in Stochastic Processes*. Academic Press.

[47] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. John Wiley.

[48] LINDLEY, D. V. (1961). Dynamic programming and decision theory. *Appl. Statist.* 10 (39-51).

[49] LIPTSER, R. S. *and* SHIRYAYEV, A. N. (1978). *Statistics of Random Processes II*. Springer-Verlag.

[50] LIPTSER, R. Sh. *and* SHIRYAYEV, A. N. (1989). *Theory of Martingales*. Kluwer Acad. Publ.

[51] MARCELLUS, R. L. (1990). A Markov renewal approach to the Poisson disorder problem. *Comm. Statist. Stochastic Models* 6 (213-228).

[52] MCKEAN, H. P., JR. (1965). Appendix: A free boundary problem for the heat equation arising from the problem in mathematical economics. *Indust. Manage. Rev.* 6 (32-39).

[53] MIKHALEVICH, V. S. (1958). A Bayes test of two hypotheses concerning the mean of a normal process. *Visnik Kiiv. Univ.* No. 1 (Ukrainian) (101-104).

[54] MYNENI, R. (1992). The pricing of the American option. *Ann. Appl. Probab.* 2 (1-23).

[55] PEDERSEN, J. L. (1997). Best bounds in Doob's maximal inequality for Bessel processes. *Research Report* No. 373, *Dept. Theoret. Statist. Aarhus* (10 pp). *J. Multivariate Anal.* 75, 2000 (36-46).

[56] PONTRYAGIN, L. S. BOLTYANSKI, V. G. GAMKRELIDZE, R. V. *and* MISCHENKO, E. F. (1962). *The Mathematical Theory of Optimal Processes*. John Wiley & Sons.

[57] REVUZ, D. *and* YOR, M. (1994). *Continuous Martingales and Brownian Motion*. (Second Edition) Springer-Verlag.

[58] ROGERS, L. C. G. *and* WILLIAMS, D. (1987). *Diffusions, Markov Processes, and Martingales; Volume 2: Itô's Calculus.* John Wiley & Sons.

[59] ROMBERG, H. F. (1972). Continuous sequential testing of a Poisson process to minimize the Bayes risk. *J. Amer. Statist. Assoc.* 67 (921-926).

[60] SHEPP, L. A. *and* SHIRYAEV, A. N. (1993). The Russian option: Reduced regret. *Ann. Appl. Probab.* 3 (631-640).

[61] SHEPP, L. A. *and* SHIRYAEV, A. N. (1994). A new look at the Russian option. *Theory Probab. Appl.* 39 (103-119).

[62] SHIRYAEV, A. N. (1961). The detection of spontaneous effects. *Soviet Math. Dokl.* 2 (740-743).

[63] SHIRYAEV, A. N. (1961). The problem of the most rapid detection of a disturbance in a stationary process. *Soviet Math. Dokl.* 2 (795-799).

[64] SHIRYAEV, A. N. (1967). Two problems of sequential analysis. *Cybernetics* 3 (63-69).

[65] SHIRYAEV, A. N. (1978). *Optimal Stopping Rules.* Springer-Verlag.

[66] SHIRYAEV, A. N. (1999). *Essentials of Stochastic Finance (Facts, Models, Theory).* World Scientific.

[67] SIEGMUND, D. (1985). *Sequential Analysis. Tests and Confidence Intervals*. Springer-Verlag.

[68] SNELL, J. L. (1952). Applications of martingale system theorems. *Trans. Amer. Math. Soc.* 73 (293-312).

[69] VAN MOERBEKE, P. (1976). On optimal stopping and free boundary problems. *Arch. Rational Mech. Anal.* 60 (101-148).

[70] WALD, A. (1947). *Sequenatial Analysis*. John Wiley and Sons, New York.

[71] WALD, A. (1950). *Statistical Decision Functions*. John Wiley and Sons, New York.

[72] WALD, A. *and* WOLFOWITZ, J. (1948). Optimum character of the sequential probability ratio test. *Ann. Math. Statistics* 19 (326-339).

[73] WALD, A. *and* WOLFOWITZ, J. (1949). Bayes solutions of sequential decision problems. *Proc. Nat. Acad. Sci. U.S.A.* 35 (99-102). *Ann. Math. Statistics* 21, 1950 (82-99).

[74] YONG, J. *and* ZHOU, X. Y. (1999). *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer-Verlag.

Goran Peskir
Department of Mathematical Sciences
University of Aarhus, Denmark
Ny Munkegade, DK-8000 Aarhus
home.imf.au.dk/goran
goran@imf.au.dk

SUMMARY

The thesis consists of twenty-two papers written by the author in the period 1994-2000. A short description of each paper with some historical remarks is given in Chapter 1. The four most illustrative papers are presented in Chapters 2-5. These chapters also contain 17 figures.

Paper [P1] (Chapter 2): The solution is found to the optimal stopping problem with payoff:

$$\sup_\tau E\Big(S_\tau - \int_0^\tau c(X_t)\, dt \Big)$$

where $S = (S_t)_{t \geq 0}$ is the maximum process associated with the one-dimensional time-homogeneous diffusion $X = (X_t)_{t \geq 0}$, the function $x \mapsto c(x)$ is positive and continuous, and the supremum is taken over all stopping times τ of X for which the integral has finite expectation. It is proved, under no extra conditions, that this problem has a solution, i.e. the payoff is finite and there is an optimal stopping time, if and only if the following *maximality principle* holds: The first-order nonlinear differential equation

$$g'(s) = \frac{\sigma^2\big(g(s)\big)\, L'\big(g(s)\big)}{2\, c\big(g(s)\big)\Big(L(s) - L\big(g(s)\big)\Big)}$$

admits a maximal solution $s \mapsto g_*(s)$ which stays strictly below the diagonal in $I\!R^2$. [In this equation $x \mapsto \sigma(x)$ is the diffusion coefficient and $x \mapsto L(x)$ the scale function of X.] In this case the following stopping time:

$$\tau_* = \inf\, \{\, t > 0 \mid X_t \leq g_*(S_t)\, \}$$

is proved optimal, and explicit formulas for the payoff are given. The result has a large number of applications, and may be viewed as the cornerstone in a general treatment of the maximum process.

Paper [P2] (Chapter 3): Let $B = (B_t)_{0 \leq t \leq 1}$ be standard Brownian motion started at zero, and let $S_t = \max_{0 \leq r \leq t} B_r$ for $0 \leq t \leq 1$. Consider the optimal stopping problem:

$$V_* = \inf_\tau E\big(B_\tau - S_1\big)^2$$

where the infimum is taken over all stopping times of B satisfying $0 \leq \tau \leq 1$. We show that the infimum is attained at the stopping time:

$$\tau_* = \inf\, \Big\{\, 0 \leq t \leq 1 \mid S_t - B_t \geq z_* \sqrt{1-t}\, \Big\}$$

where $z_* = 1.12\ldots$ is the unique root of the equation:

$$4\Phi(z_*) - 2z_*\varphi(z_*) - 3 = 0$$

with $\varphi(x) = (1/\sqrt{2\pi})e^{-x^2/2}$ and $\Phi(x) = \int_{-\infty}^x \varphi(y)\, dy$. The value V_* equals $2\Phi(z_*) - 1$. The method of proof relies upon a stochastic integral representation of S_1, time-change arguments, and the solution of a free-boundary problem.

Paper [P3] (Chapter 4): We present the explicit solution of the Bayesian and variational problem of sequential testing of two simple hypotheses about the intensity of an observed Poisson process. The method of proof consists of reducing the initial problem to a free-boundary differential-difference problem, and solving the latter by use of the principles of smooth and continuous fit. A rigorous proof of the optimality of the Wald's sequential probability ratio test in the variational formulation of the problem is obtained as a consequence of the solution of the Bayesian problem.

Paper [P4] (Chapter 5): The Poisson disorder problem seeks to determine a stopping time which is as close as possible to the (unknown) time of 'disorder' when the intensity of an observed Poisson process changes from one value to another. Partial answers to this question are known to date only in some special cases, and the main purpose of the present paper is to describe the structure of the solution in the general case. The method of proof consists of reducing the initial (optimal stopping) problem to a free-boundary differential-difference problem. The key point in the solution is reached by specifying when the principle of smooth fit breaks down and gets superseded by the principle of continuous fit. This can be done in probabilistic terms (by describing the sample path behaviour of the a posteriori probability process) and in analytic terms (via the existence of a singularity point of the free-boundary equation).

DANSK RESUMÉ

Afhandlingen består af 22 artikler skrevet af forfatteren i perioden 1994-2000. I kapitel 1 er der en kort beskrivelse af hver artikel med historisk baggrund. De 4 mest illustrative artikler findes i kapitel 2-5. Disse kapitler indeholder også 17 figurer.

Artikel [P1] (Kapitel 2): Løsningen til det optimale stoptidsproblem er fundet med 'payoff':

$$\sup_{\tau} E\bigg(S_\tau - \int_0^\tau c(X_t)\, dt \bigg)$$

hvor $S = (S_t)_{t \geq 0}$ er den maksimale proces forbundet med en endimensional tidshomogen diffusion $X = (X_t)_{t \geq 0}$, funktionen $x \mapsto c(x)$ er positiv og kontinuert, og supremaet tages over alle stoptider τ for hvilket integralet har endelig middelværdi. Det bevises, uden ekstra betingelser, at dette problem har en løsning, dvs. den tilhørende 'payoff' er endelig og der findes en optimal stoptid, hvis og kun hvis det følgende *maksimalitet princip* holder: Den først-orden ikke-linære differentiale ligning

$$g'(s) = \frac{\sigma^2\big(g(s)\big)\, L'\big(g(s)\big)}{2\, c\big(g(s)\big)\Big(L(s) - L\big(g(s)\big)\Big)}$$

har en maksimal løsning $s \mapsto g_*(s)$, som forbliver under diagonalen i $I\!R^2$. [I denne ligning er $x \mapsto \sigma(x)$ diffusionskoefficienten og $x \mapsto L(x)$ er skalafunktionen for X.] I dette tilfælde bevises, at følgende stoptid:

$$\tau_* = \inf \{\, t > 0 \mid X_t \leq g_*(S_t) \,\}$$

er optimal, og eksplicitte formler for 'payoff' er givet. Resultatet har en bred vifte af anvendelser og kan blive set som en hjørnesten i en general behandling af den maksimale proces.

Artikel [P2] (Kapitel 3): Lad $B = (B_t)_{0 \leq t \leq 1}$ være standard Brownske bevægelse begyndende i nul og lad $S_t = \max_{0 \leq r \leq t} B_r$ for $0 \leq t \leq 1$. Betragt det optimale stoptidsproblem:

$$V_* = \inf_{\tau} E\big(B_\tau - S_1\big)^2$$

hvor infimaet tages over alle stoptider for B , som tilfredsstiller $0 \leq \tau \leq 1$. Vi viser at infimaet antages for stoptiden:

$$\tau_* = \inf \big\{\, 0 \leq t \leq 1 \mid S_t - B_t \geq z_* \sqrt{1-t} \,\big\}$$

hvor $z_* = 1.12\ldots$ er entydige rod i ligningen:

$$4\Phi(z_*) - 2z_*\varphi(z_*) - 3 = 0$$

med $\varphi(x) = (1/\sqrt{2\pi})e^{-x^2/2}$ og $\Phi(x) = \int_{-\infty}^x \varphi(y)\, dy$. Værdien V_* er lig med $2\Phi(z_*) - 1$. Bevismetoden afhænger af en stokastisk integral beskrivelse af S_1 , tids-skiftsargumenter, og løsningen af et ' free-boundary' problem.

Artikel [P3] (Kapitel 4): Vi præsenterer den eksplicitte løsning for den Bayesianske og den variationelle problemstilling, der opstår ved at betragte løbende test for to simple hypoteser vedrørende intensiteten af en observeret Poisson proces. Bevismetoden består i at reducere det oprindelig problem til et 'free-boundary' differens-differential problem. Det sidstnævnte løses ved hjælp af 'smooth' og 'continuous fit' principper. Et rigoristisk bevis for den Wald's sekventielle kvotient test i den variationelle problemstilling opnåes som en konsekvens af løsningen på det Bayesianske problem.

Artikel [P4] (Kapitel 5): Det såkaldte 'Poisson disorder' problem søger at fastsætte en stoptid som er så tæt som muligt på det (ukendte) tidspunkt for 'uorden', dvs. intensiteten af et observeret Poisson proces skifter fra en værdi til en anden. Delvise svar på dette spørgsmål er i dag kendt kun i nogle specielle tilfælde, og hovedformålet med denne afhandling er at beskrive strukturen for løsningen i det generelle tilfælde. Bevismetoden består i at reducere det oprindelig (optimale stoptids) problem til et 'free-boundary' differens-differential problem. Kardinalpunktet for løsningen består i at specificere, hvornår princippet for 'smooth fit' bryder sammen og bliver erstattet af princippet for 'continuous fit'. Dette bliver gjort i sandsynlighedsteoretiske termer (ved at beskrive udfaldsfunktionens opførsel for en 'a posteriori' sandsynlighedsproces) og i analytiske termer (via eksistensen af et singularitetspunkt for den tilhørende 'free-boundary' ligning).