# The High Contact Principle as a Sufficiency Condition for Optimal Stopping

Kjell Arne Brekke* and Bernt Øksendal †‡

January 29, 1990

## 1 Introduction

The "high contact" principle was first introduced by Samuelson (1965). He only gave a heuristic argument for the condition. In McKean's mathematical appendix to Samuelson's paper (McKean (1965)), a rigorous proof for the necessity of the condition was given for the case of a linear reward and a geometrical Brownian motion process, i.e. he proved that any solution of the optimal stopping problem has to satisfy the high contact condition. As there only exists one function satisfying the high contact condition in this situation, Samuelson's proposal is the only possible optimal solution.

In Shiryayev (1978, Theorem 3.17) the condition appears as a necessary condition for one dimensional processes (but not under the name "high contact"). Multidimensional versions of the theorem are given in Friedman (1976) and Bensoussan and Lions (1982). These theorems are derived from variational inequalities, and the assumptions in the theorems are too strong to apply to most economic applications. (See section 2.)

The high contact principle is essentially a first order condition in the optimization of the stopping time (see Merton (1973, footnote 60), and for a rigorous further development of the same idea Øksendal (1990)). To derive the "second order conditions" turns out to be easy, and we will in this paper prove, under weak conditions, that a solution proposal to an optimal stopping problem satisfying the high contact principle, is in fact an optimal solution to the problem. In this case we do not have to prove that there is only one solution satisfying the high contact principle, and the existence of the optimal solution is a part of the conclusion.

Some results from stochastic analysis will be used without reference, these results can be found in Øksendal (1989).

*Research Department, Central Bureau of Statistics, Box 8131 Dep, N-0033 Oslo1, Norway.
†Institute of Mathematics, University of Oslo, Box 1053 Blindern, N-0316 Oslo 3, Norway.
‡We thank Trond Olsen and Henrik H. Martens for comments to the first version of this paper.

## 2 The problem

Let

$$dX_t = b(X_t)dt + \sigma(X_t)dB_t \tag{1}$$

be an n-dimensional Itô diffusion, where $b : \mathbb{R}^n \mapsto \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \mapsto \mathbb{R}^{n \times m}$ are Lipschitz continuous functions with at most linear growth. Let $g$ be a (real) bounded continuous function on $\mathbb{R}^n$. The optimal stopping problem is the problem of finding:

$$g^*(x) = \sup_\tau E^x[g(X_\tau)] \tag{2}$$

the sup being taken over all $\mathcal{F}_t$-stopping times $\tau$, where $\mathcal{F}_t$ is the $\sigma$-algebra generated by $B_s$, $s \leq t$. Here $E^x$ denotes the expectation w.r.t. the law $P^x$ of $\{X_t\}$ given $X_0 = x$. The *optimal stopping time* corresponding to $g^*$ is denoted $\tau^*$.

If we know $g^*$ it is easy to find $\tau^*$: It is obviously optimal to stop if $g(X_t) \geq g^*(X_t)$ since we then achieve the optimal benefit, while if $g(X_t) < g^*(X_t)$ it is not optimal to stop, since this would give less than the optimum. The set

$$D = \{x : g^*(x) > g(x)\}$$

is called the continuation region. Obviously, for $x \in D$,

$$g^*(x) = E^x\{g(X_{\tau_D})\}$$

where $\tau_D = \inf\{s > t : X_s \notin D\}$.

If the continuation region is known, then the problem of finding $g^*$ can be transformed into a Dirichlet problem. Define the operator

$$L = \sum_i b_i(x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j} a_{i,j}(x)\frac{\partial^2}{\partial x_i \partial x_j} \tag{3}$$

where $a = \sigma\sigma^T$, and where $b$ and $\sigma$ are as in (1). Then it is known that $g^*$ solves the Dirichlet problem:

$$\begin{aligned} &(Lg^*)(x) = 0 && \text{for } x \in D \\ &\lim_{x \to y} g^*(x) = g(y) && \text{for all regular } y \in \partial D. \end{aligned} \tag{4}$$

($y \in \partial D$ is called regular if $\tau_D = 0$ a.s. $P^y$.)

But $D$ is unknown, so this is a free boundary problem. Therefore an additional boundary condition is needed to identify the boundary. This is why the "high contact" principle is important. The principle states that

$$\nabla g^* = \nabla g \qquad \text{on } \partial D, \tag{5}$$

which gives us the extra boundary condition.

In a footnote Merton (1969) derives the high contact principle as a first order condition. Suppose $X_t$ is one dimensional. Let $D_c = \{x : x < c\}$, $\tau_c = \tau_{D_c}$, and

$$f(x, c) = E^x[g(X_{\tau_c})] \tag{6}$$

Note that if $c$ is regular for $D_c$ then $f(c,c) = g(c)$, and hence if $f'_c$ exists, we have

$$f'_x(c,c) + f'_c(c,c) = g'(c) \tag{7}$$

Suppose

$$g^*(x) = f(x,c^*) = \max_c f(x,c),$$

then the high contact condition $f'_x(c^*,c^*) = g'(c^*)$ is a direct consequence of the first order condition $f'_c(x,c^*) = 0$.

To make this argument rigorous it is necessary to verify that:

(i) $D$ is of the form $D_c$, and $c$ is a regular point for $D_c$.

(ii) $c \mapsto f(x,c) \in C^1$

And if we want to use this high contact principle to prove that a proposed solution is optimal we must show that:

(iii) The candidate satisfying high contact is unique

(iv) There exists an optimal solution

In the multidimensional case, which is the most relevant one for economic applications, it is necessary to have smoothness assumptions on the *optimal* continuation regions as well. In Øksendal (1990) a property analogous to (ii) is verified under the assumption that $L$ is elliptic. Since the argument only proves the *necessity* of high contact, we must also verify (iii) and (iv).

The purpose of this paper is to show that – under certain conditions – the high contact property is also *sufficient* for the solution of the optimal stopping problem. More precisely, we show that if there exists an open set $D \subset \mathbb{R}^n$ with $C^1$ – boundary and a function $h$ on $D$ such that

$$h \geq g \qquad \text{on } D \tag{8}$$

$$Lg \leq 0 \qquad \text{outside the closure } \overline{D} \tag{9}$$

("the second order condition") and such that $(D,h)$ solves the free boundary problem

$$Lh = 0 \qquad \text{on } D \tag{10}$$

$$h = g \qquad \text{on } \partial D \tag{11}$$

$$\nabla h = \nabla g \qquad \text{on } \partial D \tag{12}$$

then in fact $h = g^*$ on $D$.

To achieve this result the basic idea is the following. Extend $h$ to $\mathbb{R}^n$ by setting $h = g$ outside $D$. We know that $g^*$ is the least superharmonic majorant of $g$. We also know that $h \leq g^*$ since $h$ is what we get from using $\tau = \tau_D$. It only remains to show that $h$ is $X_t$–superharmonic, i.e:

$$h(x) \geq E^x[h(X_\tau)] \tag{13}$$

for all stopping times $\tau$, and all $x$. If $h \in C^2$, then this is equivalent to $Lh \leq 0$. This follows from Dynkin's formula:

$$h(x) = E^x[h(X_\tau) - \int_0^\tau Lh(X_t)dt] \tag{14}$$

By construction $Lh = 0$ in $D$, and by assumption $Lh \leq 0$ outside the closure $\overline{D}$. Unfortunately we generally only have $h \in C^1(\partial D)$ (this is the high contact principle). If we can approximate $h$ with $\hat{h} \in C^2$ such that $|h - \hat{h}| < \epsilon$ and $L\hat{h} \leq \epsilon$, for an arbitrary small $\epsilon$, this will do. In Brekke (1989, Appendix C) this idea is used to prove the sufficiency of the high contact in an essentially one-dimensional case.

To extend this result to the multidimensional case, it turns out to be more convenient to generalize the Dynkin formula, using a Green function. This is done in Section 3, but first we will consider the relevance of the theorem in economic applications.

# 3 High contact in economics

As we have pointed out in the introduction, the existing high contact theorems are either one-dimensional or make very strong assumptions. Let us consider a typical problem in the economics of exhaustible resources, the problems will be similar in other economic applications. Let

$$dP_t = \alpha P_t dt + \beta P_t dB_t \tag{15}$$

be the price of the resource, where $\alpha$ and $\beta$ are constants. Consider the stopping problem

$$\gamma^*(t,p) = \sup_{\tau} E^{t,p}\{\int_t^{\tau} f(P_s)e^{-rs}ds + g(P_\tau)e^{-r\tau}\} \tag{16}$$

where $f$ and $g$ are given bounded continuous functions. If the problem is to find the optimal time to stop production, then $f$ is the profit, and $g$ is the abandonment cost. When the problem is to find the optimal time to start a project, $g(P)$ is the net present value of the field started at price $P$, and $f = 0$. Note that (16) is not of the form (2), because time and the integral of $f$ until $\tau$ are introduced in the reward function. The problem can, however, be brought into the form of (2) at the cost of increasing the dimension of the process. Let $Y_t = (t, P_t, \Theta_t)$ where

$$\Theta_t = \theta + \int_0^t f(P_s)e^{-rs}ds \tag{17}$$

is the "profit" earned until $t$, then

$$g^*(t,p) + \theta = sup_\tau E^{t,p,\theta}(g(\tau, P_\tau) + \Theta_\tau) \tag{18}$$

A high-contact principle applying to this problem must allow for processes of dimension three (or two if $f = 0$). In other words, none of the one–dimensional theorems apply to economic problems where discounting is relevant.

Multi–dimensional results derived from regularity results for variational inequalities make assumptions that exclude the geometrical Brownian motion (15). A theorem in Friedman (1976) requires e.g. that $L$ is uniformly elliptic and that $a(x) = \sigma\sigma^T$

is bounded, which excludes the geometrical Brownian motion where $a(p) = \beta^2 p^2$. Furthermore he assumes that

$$b_i(x) = \sum_{j=1}^{n} \frac{\partial a_{ij}(x)}{\partial x_j}$$

which in the case of a geometrical Brownian motion means $\alpha = 2\beta^2$. These assumptions are too strong for most economic applications.

Another multi–dimensional result is given in Øksendal (1989), who proves the necessity of the high contact principle under the assumption that $L$ is locally elliptic, but he has to make smoothness assumptions on the form of the boundary as well.

All these results are necessary conditions. Here we will establish that high contact is sufficient for a function to be identified as a solution of the optimal stopping problem. Moreover, we can relax the assumptions to cover most economic applications, e.g. the time–space geometrical Brownian motion mentioned above.

# 4 A generalized Dynkin formula

This section uses some advanced mathematical results and methods, and hence it may be difficult to read. Lemma 1 and the following remark contains the results that are needed in the later sections. The reader that is not interested in the details, may skip the remaining part of the section.

Let $X_t$ be as defined in (1) and $L$ the corresponding operator (3). Then $X_t$ has a generator which coincides with $L$ on the smooth functions. Let $L^q_{\text{loc}}(dx)$, $q \geq 1$ be the set of all functions $f$ such that $|f|^q$ is locally integrable with respect to the Lebesgue measure $dx$, and $C^2_b(A)$ is the set of functions with bounded continuous derivatives of second order in the set $A$.

If $V \subset \mathbb{R}^n$ is a bounded domain such that $E^\xi[\tau_V] < \infty$ for all $\xi$ then we can define the *Green measure* of $X$ (with respect to $V$), $\mathcal{G}(\xi, \cdot)$, by

$$\mathcal{G}(\xi, \phi) = \mathcal{G}_V(\xi, \phi) = E^\xi[\int_0^{\tau_V} \phi(X_t)dt] \qquad \phi \in C(\overline{V}) \tag{19}$$

If the measure $\mathcal{G}(\xi, dx)$ has the form

$$\mathcal{G}(\xi, dx) = G(\xi, x)dx \qquad (dx \text{ is the Lebesgue measure}) \tag{20}$$

then we say that $X$ has a *Green function* $G(\xi, x)$ (in $V$). Note that in this case we can write

$$E^\xi[\int_0^{\tau_V} \phi(X_t)dt] = \int_V \phi(x)G(\xi, x)dx \tag{21}$$

A sufficient condition that $X$ has a Green function in any bounded domain $V$, is that $X$ (or, more precisely, the generator $L$ of $X$) is *uniformly elliptic* in $V$, i.e. that there exists $\lambda > 0$ such that

$$z^T a(x)z \geq \lambda |z|^2 \qquad \text{for all } x \in V, z \in \mathbb{R}^n \tag{22}$$

In fact, if $X$ is uniformly elliptic in $V$ and $n \geq 3$ then by a result of Littman, Stampacchia and Weinberger (1963) there exists, for any compact $H$ in $V$, a constant $C = C(H) < \infty$ such that

$$G^X(\xi, x) \leq C \cdot |\xi - x|^{2-n} \qquad \text{for all } \xi, x \in H. \tag{23}$$

Using polar coordinatesin $\mathbb{R}^n$ we see that this implies that

$$\int_H G(\xi, x)^q dx = C_1 \int_0^1 r^{-(n-2)q} r^{n-1} dr = C_1 \int_0^1 r^{n-1-q(n-2)} dr \tag{24}$$

which is finite if $q < 1 + \frac{n}{2}$.

We summarize:

If $X$ is uniformly elliptic in $V$ then $X$ has a Green function $G(\xi, x)$ satisfying

$$G(\xi, x) \in L^q_{loc}(dx) \qquad \text{for } q < 1 + \frac{1}{n} \tag{25}$$

We are also interested in processes of the form

$$dX_t = \begin{bmatrix} dK_t \\ dY_t \end{bmatrix} = \begin{bmatrix} \rho(K_t) \\ \mu(Y_t) \end{bmatrix} dt + \begin{bmatrix} 0 \\ \nu(Y_t) \end{bmatrix} dB_t \tag{26}$$

with $K_t \in \mathbb{R}^m$ and $Y_t \in \mathbb{R}^d$.

Assume for simplicity that $V = M \times N$ where $M \subset \mathbb{R}^m$, $N \subset \mathbb{R}^d$ and assume that $Y$ is uniformly elliptic in $N$. Then it is well known that $Y$ has a *transition function* $p_t(\eta, y)$ in $N$, in the sense that

$$E^\eta[\psi(Y_t) \mathcal{X}_{t<\tau_N}] = \int_N \psi(y) p_t(\eta, y) dy, \tag{27}$$

where $dy$ denotes Lebesgue measure (in $\mathbb{R}^d$).

In fact, by a result of Aronson (1967) there exist constants $C < \infty$, $\alpha > 0$ such that

$$p_t(\eta, y) \leq C \cdot t^{-d/2} \exp(-\frac{\alpha|\eta - y|^2}{t}) \tag{28}$$

for all $t$, and all $x, y \in N$.

Using the transition function of $Y$ we can describe the Green measure $\mathcal{G}_X(\xi, \cdot)$, for $X$ (given by (26)) as follows:

Suppose $\phi(x) = \phi_1(k)\phi_2(y)$ where $x = (k, y)$. Then

$$\begin{aligned} \mathcal{G}_X(\xi, \phi) &= E^\xi\left[\int_0^\infty \phi_1(K_t)\phi_2(Y_t) \cdot \mathcal{X}_{t<\tau_V} dt\right] \\ &= \int_0^\infty \left(\int_N \phi_1(K_t) \cdot \phi_2(y) p_t(\eta, y) dy\right) \mathcal{X}_{K_t \in M} dt, \quad \xi = (k_0, \eta). \end{aligned} \tag{29}$$

Since a general $\phi(k, y)$ can be approximated by a sum of such products, we conclude that for general $\phi(k, y)$ we have

$$\mathcal{G}_X(\xi, \phi) = \int_0^\infty \left(\int_N \phi(K_t, y) p_t(\eta, y) dy\right) \mathcal{X}_{K_t \in M} dt \tag{30}$$

with $p_t(\eta, y)$ satisfying (28), provided that $X_t = (K_t, Y_t)$ where $Y_t$ is uniformly elliptic.

The conclusions (25) and (30) will be needed in the proof of the next result.

**Lemma 1 (Generalized Dynkin formula)**

*Let $U, V$ be bounded domains with $C^1$ borders in $\mathbb{R}^n$ and let $\Gamma$ denote the boundary of $U$. Suppose*

$$\phi \in C_b^2(V \setminus \Gamma) \cap C^1(\overline{V}) \tag{31}$$

*and that $X$ satisfies one of the following two conditions:*

1. *$X_t$ is uniformly elliptic in $V$*
2. *$X_t = (K_t, Y_t)$ as in (26), with $Y_t$ uniformly elliptic in $N$ (where $V \subset M \times N$) and for each $k \in M$ the set*

$$\Gamma_k = \{y \in N : (k, y) \in \Gamma\} \subset \mathbb{R}^d \tag{32}$$

*has zero d-dimensional Lebesgue measure.*

*Then*

$$E^{\xi}[\phi(X_{\tau_V})] = \phi(\xi) + E^{\xi}[\int_0^{\tau_V} L\phi(X_t)\,dt] \tag{33}$$

*where $L\phi(x)$ is the function defined for all $x \in V \setminus \Gamma$ by pointwise differentiation according to (3).*

**Remark:** We will use (33) in the form of (44) (in case 1) or (46) (in case 2). In both cases, the expectation is transformed to an integral with respect to the Lebesgue measure. Thus, since the boundary $\Gamma$ has Lebesgue measure zero, we do not have to define $L\phi$ on the boundary.

**Proof**

First assume that condition 1 holds.

Let $D_{jk}\phi$ denote the distributional double derivative of $\phi$ with respect to $x_j$ and $x_k$ and let $\frac{\partial^2\phi}{\partial x_j \partial x_k}$ denote the pointwise double derivative which is defined everywhere outside $\Gamma$ and hence almost everywhere ($dx$). We claim that

$$D_{jk}\phi = \frac{\partial^2 \phi}{\partial x_j \partial x_k} \qquad \text{for } 1 \le j, k \le n \tag{34}$$

To establish (34) we put $V_1 = V \cap U$, $V_2 = V \setminus \overline{U}$ and choose $u \in C_0^\infty(V)$. By integration by parts we have

$$\int_{V_i} \frac{\partial^2 \phi}{\partial x_j \partial x_k} u\,dx = \int_{\partial V_i} \frac{\partial \phi}{\partial x_k} n_{ij} u\,ds - \int_{V_i} \frac{\partial \phi}{\partial x_k}\frac{\partial u}{\partial x_j} dx, \tag{35}$$

where $n_{ij}$ is component $j$ of the outer unit normal $\vec{n}_i$ from $V_i$. Another integration by parts leads to

$$\int_{V_i} \frac{\partial \phi}{\partial x_k}\frac{\partial u}{\partial x_j} dx = \int_{\partial V_i} \phi \frac{\partial u}{\partial x_j} n_{ik}\,ds - \int_{V_i} \phi \frac{\partial^2 u}{\partial x_j \partial x_k} dx \tag{36}$$

Combining (35), (36) and adding for $i = 1,2$ gives:

$$\int_V \frac{\partial^2 \phi}{\partial x_j \partial x_k} u dx = \int_V \phi \frac{\partial^2 u}{\partial x_j \partial x_k} dx \tag{37}$$

which proves the claim (34).

For $m = 0,1,2,...$ and $1 \leq p < \infty$ define the spaces

$$W^{m,p} = \left\{ u : \begin{array}{l} u \in L^p(V); D^\alpha u \in L^p(V) \text{ for all multi-indices} \\ \alpha = (\alpha_1, ..., \alpha_k) \text{ with } |\alpha| = \alpha_1 + ... + \alpha_k \leq m \end{array} \right\} \tag{38}$$

equipped with the norm

$$\|u\|_{m,p} = \left( \sum_{|\alpha| \leq m} \|D^\alpha u\|_p^p \right)^{1/p} \tag{39}$$

and let $H^{m,p}$ denote the closure of $C^m$ in this norm. Then a famous result of Meyers & Serrin ( see e..g. Adams (1975), Th. 3.16) states that

$$H^{m,p} = W^{m,p} \tag{40}$$

For all $p < \infty$ we have $\phi \in W^{2,p}$, by (31) and (34). So by (40) there exists a sequence $\{\phi_k\} \in C^2$ such that

$$\|\phi_k - \phi\|_{2,p} \to 0 \quad \text{as } k \to \infty \tag{41}$$

If $p > \frac{n}{2}$ the Sobolev inequality combined with (41) gives that

$$\phi_k \to \phi \text{ uniformly on } V \tag{42}$$

Since $\phi_k \in C^2$, we know that Dynkin's formula holds for $\phi_k$, i.e.

$$E^\xi[\phi_k(X_{\tau_V})] = \phi_k(\xi) + \int_V L\phi_k(x) G(\xi, x) dx \tag{43}$$

Choose $p > \frac{n}{2} + 1$, then if $\frac{1}{p} + \frac{1}{q} = 1$ we have $q < 1 + \frac{2}{n}$, so for such a value of $p$ we can combine (41), (42) and (43) to conclude that:

$$E^\xi[\phi(X_{\tau_V})] = lim_{k\to\infty} E^\xi[\phi_k(X_{\tau_V})] = \phi(\xi) + \int_V L\phi(x) G(\xi, x) dx \tag{44}$$

because by Hölder's inequality

$$|\int_V (L\phi_k - L\phi)(x) G(\xi, x) dx| \leq \|L\phi_k - L\phi\|_p \|G(\xi, \cdot)\|_q \to 0 \tag{45}$$

as $k \to \infty$ by (25). That proves (33).

Next assume that the second condition holds. We proceed as in the first case up to (43), so that we have, using (30), for each $k$

$$E^\xi[\phi_k(X_{\tau_V})] = \phi_k(\xi) + \int_0^\infty \left( \int_N L\phi_k(K_t, y) p_t(\eta, y) dy \right) \mathcal{X}_{K_t \in M} dt \tag{46}$$

Choose $1 < q < \infty$ (to be determined later) and apply Hölder's inequality for each $t$:

$$\begin{aligned} & \int_W |L\phi_k(K_t, y) - L\phi(K_t, y)| p_t(\eta, y) dy \\ \le & \left(\int_W |L\phi_k - L\phi|^p dy\right)^{1/p} \left(\int_W p_t^q(\eta, y) dy\right)^{1/q} \quad \text{where } \tfrac{1}{p} + \tfrac{1}{q} = 1 \end{aligned} \tag{47}$$

By the estimate (28) we get, using the substitution $u = \frac{y}{\sqrt{t}}$:

$$\int_W p_t^q(\eta, y) dy \le C_1 \cdot t^{-qd/2} \int_{\mathbb{R}^d} \exp(-\alpha q |u|^2) t^{d/2} du \le C_2 \cdot t^{-(q-1)d/2} \tag{48}$$

and this is locally $t$-integrable near 0 if $q < 1 + \frac{2}{d}$. (Note that we used the Hölder inequality for each $t$, so we need a $t$-uniform approximation, but this follows from the result of Meyers and Serrin which implies that there exist a constant $C_3$ such that

$$\|\phi_k - \phi\|_{2,p} < C_3 \|\phi\|_{2,p}$$

for all $k$ and all $t$, where $\|\phi\|_{2,p}$ means the norm of the function $y \mapsto \phi(K_t, y)$. )
Therefore, if we choose $1 < q < 1 + \frac{2}{d}$ and $p$ such that $\frac{1}{p} + \frac{1}{q} = 1$ we obtain that

$$E^{\xi}\left[\int_0^{\tau_V} L\phi_k(X_t) dt\right] \to E^{\xi}\left[\int_0^{\tau_V} L\phi(X_t) dt\right] \tag{49}$$

as before, thereby completing the proof of Lemma 1. ■

# 5 The sufficiency of high contact

We now apply this to the optimal stopping problem (2):

$$g^*(x) = \sup_{\tau} E^x[g(X_\tau)] = E^x[g(X_{\tau^*})] \tag{50}$$

**Theorem 1** *(Sufficiency of high contact for the optimal stopping problem)*
*Suppose $W \subset \mathbb{R}^n$ is an open set such that $X_t \in W$ for all $t$ if $X_0 \in W$. Let $g \in C^1(W)$. Suppose we can find an open set $D \subset W$ with $C^1$ boundary such that $\tau_D < \infty$ a.s.. Assume furthermore that $X_t$ satisfies either condition 1. or 2. of Lemma 1 for $U = D$ and for every sufficiently small open ball $V$ centred on the boundary $\partial D \cup W$, and suppose we can find a function $h$ on $\overline{D}$ such that $h \in C^1(\overline{D}) \cap C^2(D)$, $h \ge g$ on $D$, $g \in C^2(W \setminus \overline{D})$ and $Lg \le 0$ in $W \setminus \overline{D}$, and such that $(D, h)$ solves the free boundary problem:*

$$\begin{array}{llcll} (i) & Lh(x) & = & 0 & \text{for } x \in D \\ (ii) & h(x) & = & g(x) & \text{for } x \in \partial D \\ (iii) & \nabla_x h(x) & = & \nabla_x g(x) & \text{for } x \in \partial D \cap W \\ & & & & \text{if } X_t \text{ satisfies 1.} \\ (iii)' & \nabla_y h(x) & = & \nabla_y g(x) & \text{for } x \in \partial D \cap W \\ & & & & \text{if } X_t \text{ satisfies 2.} \end{array} \tag{51}$$

*(where $x = (k, y)$ if $X_t$ satisfies 2.)*

*Extend $h$ to all of $W$ by putting $h = g$ outside $D$. Then $h$ solves the optimal stopping problem (50), i.e.:*

$$h(x) = g^*(x) = \sup_\tau E^x[g(X_\tau)] \tag{52}$$

*and thus $\tau^* = \tau_D$ is an optimal stopping time.*

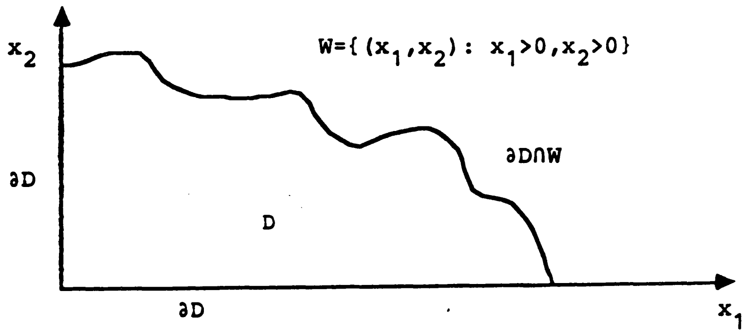


Figure 1: Illustration of boundaries

**Remark:** Note that in (51), $\partial D$ is the boundary of the set $D$ regarded as a subset in $\mathbb{R}^n$, not relative to $W$. Thus there may be parts of $\partial D$ not belonging to $W$, as illustrated in Figure 1, where $W$ is the positive orthant of $\mathbb{R}^2$. Note also that (51) *(iii)* or *(iii)'* is the high contact condition.

**Proof:** First we note that by (51) (i), (ii) we have

$$h(x) = E^x[g(X_{\tau_D})] \tag{53}$$

hence $h \leq g^*$. To prove the opposite inequality, it suffices to show that $h$ is $X_t$–superharmonic since we have assumed that $h \geq g$ and we know that $g^*$ coincides with the least $X_t$–superharmonic majorant of $g$. For this it is enough to show that $h$ is *locally* $X_t$–superharmonic. (Dynkin (1965), p.22) Since $Lh \leq 0$ outside $\partial D$, $h$ is clearly $X_t$–superharmonic there. So it remains to show that $h$ is $X_t$–superharmonic on $\partial D$:

Fix $\xi \in \partial D$ and let $V$ be a ball centred at $\xi$. By Lemma 1 we get

$$E^\xi[h(X_\tau)] = h(\xi) + \int_V Lh(x) G(\xi, x) dx \leq h(\xi), \tag{54}$$

where $\tau = \tau_V$. Thus $h$ is locally $X_t$–superharmonic everywhere and the proof is complete. ■

This result also applies to the apparently more general problem

$$\gamma^*(x) = sup_\tau E^x[\int_0^\tau f(X_s) ds + g(X_\tau)] \tag{55}$$

where $f(x)$ is a given bounded continuous function. In this case we consider the process $Z_t$ given by:

$$dZ_t = \begin{bmatrix} dX_t \\ d\Theta_t \end{bmatrix} = \begin{bmatrix} b(X_t) \\ f(X_t) \end{bmatrix} dt + \begin{bmatrix} \sigma(X_t) & | & 0 \\ 0 & | & 1 \end{bmatrix} d\tilde{B}_t, \qquad Z_0 = (x, \theta) \tag{56}$$

where $\tilde{B}_t = (B_1(t), ..., B_{n+1}(t))$ is an $(n+1)$–dimensional Brownian motion. Consider the problem

$$k^*(z) = sup_\tau E^z[k(Z_\tau)], \tag{57}$$

where $k(z) = k(x, \theta) = g(x) + \theta$.

For $E[\tau] < \infty$ we have $E^0[B_{n+1}(\tau)] = 0$, and since it is enough to take the sup over such stopping times we have

$$\begin{aligned} k^*(x,\theta) &= sup_\tau E^x[\theta + \int_0^\tau f(X_s)ds + B_{n+1}(\tau) + g(X_\tau)] \\ &= \theta + sup_\tau E^x[\int_0^\tau f(X_s)ds + g(X_\tau)] \\ &= \theta + \gamma^*(x) \end{aligned} \tag{58}$$

Therefore, if $k^*$ solves the problem (57), then $\gamma^* = k^* - \theta$ solves (55). This gives the following conclusion.

**Theorem 2** *Suppose $W \subset I\!R^n$ is an open set such that $X_t \in W$ for all $t$ if $X_0 \in W$. Let $g \in C^1(W)$. Suppose we can find an open set $D \subset W$ with $C^1$ boundary such that $\tau_D < \infty$ a.s.. Assume furthermore that $X_t$ satisfies either condition 1. or 2. of Lemma 1 for $U = D$ and for every sufficiently small open ball $V$ centred on the boundary $\partial D \cup W$, and a function $h$ on $\overline{D}$ such that $h \in C^1(\overline{D}) \cap C^2(D)$, $h \geq g$ on $D$, $g \in C^2(W \setminus \overline{D})$ and $Lg \leq -f$ in $W \setminus \overline{D}$, and such that $(D, h)$ solves the free boundary problem:*

| | | | | |
|---|---|---|---|---|
| (a) | $Lh(x)$ | $=$ | $-f(x)$ | *for* $x \in D$ |
| (b) | $h(x)$ | $=$ | $g(x)$ | *for* $x \in \partial D$ |
| (c) | $\nabla_x h(x)$ | $=$ | $\nabla_x g(x)$ | *for* $x \in \partial D \cap W$ *if* $X_t$ *satisfies 1.* |
| (c)′ | $\nabla_y h(x)$ | $=$ | $\nabla_y g(x)$ | *for* $x \in \partial D \cap W$ *if* $X_t$ *satisfies 2.* |

(59)

*(where $x = (k, y)$ if $X_t$ satisfies 2.)*

*Suppose furthermore that for all $x$, we can find $p > 1$, such that:*

$$\inf_{\omega \in \Omega, T > 0} \int_0^T |f(X_s(\omega))| ds > -\infty. \tag{60}$$

*where $\Omega$ is the measure space on which $\tilde{B}_t$ is defined.*

*Extend $h$ to all of $W$ by putting $h = g$ outside $D$. Then $h$ solves the optimal stopping problem (55), i.e.:*

$$h(x) = \gamma^*(x) = \sup_\tau E^x[\int_0^\tau f(X_s)ds + g(X_\tau)] \tag{61}$$

*and thus $\tau^* = \tau_D$ is an optimal stopping time.*

**Proof:** Define $H(x,\theta) = h(x) + \theta$, $k(x,\theta) = g(x) + \theta$.
Let

$$\mathcal{L}u(x,\theta) = L_x u + f(x)\frac{\partial u}{\partial \theta} + \frac{1}{2}\frac{\partial^2 u}{\partial \theta^2}, \quad u \in C^\infty \tag{62}$$

denote the generator of the process $Z_t$.

Then by (a), (b) and (c) we have

$$\mathcal{L}H(x,\theta) = Lh(x) + f(x) = 0 \text{ if } x \in D \tag{63}$$

$$H(x,\theta) = k(x,\theta) \quad \text{if } x \in \partial D \tag{64}$$

and

$$\begin{cases} \nabla_x H(x,\theta) &= \nabla_x k(x,\theta) \\ \frac{\partial H}{\partial \theta}(x,\theta) &= \frac{\partial k}{\partial \theta}(x,\theta) \end{cases} \quad \text{if } x \in \partial D \tag{65}$$

By Theorem 1 applied to $h = H$, $g = k$ and the process $Z_t$, and by (58) we conclude that

$$H(x,\theta) = sup_\tau E^{x,\theta}[k(Z_\tau)] = \theta + \gamma^*(x) \tag{66}$$

i.e. that $h = \gamma^*$ as claimed.

Note that $k = \theta + g$ is not bounded, hence Theorem 1 does not apply directly. By an inspection of the proof, however, we find that the condition (60) is sufficient. ■

**Remark:** If $f$ is of the form $f(x) = F(y)e^{-rt}$, where $F(y) \geq -M > -\infty$, then

$$\int_0^T f(X_t(\omega))dt \geq -\frac{M}{r} > -\infty \tag{67}$$

hence (60) is satisfied. This applies to many economic problems.

# 6 An application: Starting and stopping of resource extraction

The starting and stopping of a mine or a field was studied in a seminal paper by Brennan and Schwartz (1982). For ease of analysis we will simplify their model considerably, and disregard the discussion of taxes and convenience yield. The results in this section is also similar to the entry and exit model of Dixit (1989), where the only difference is that his model includes no resource extraction.

We will formulate the model as two simultaneous stopping problems. The use of the high contact principle in this model illustrates the use of both theorems. Furthermore, we point at an unresolved problem in using optimal stopping theories to solve sequential stopping problems.

Suppose that the price process is a geometrical Brownian motion:

$$dP_t = \alpha P_t dt + \beta P_t dB_t \tag{68}$$

where $\alpha$, $\beta$ are constants. The stock of remaining reserves in the field is denoted by $Q_t$. If the field is open, extraction is proportional to remaining reserves. Hence

$$dQ_t = -\mathcal{X}\lambda Q_t dt \tag{69}$$

where

$$\mathcal{X} = \begin{cases} 1 & \text{if the field is open} \\ 0 & \text{if the field is closed} \end{cases} \tag{70}$$

When the field is open, a rental cost $K$ is the only operating cost. Thus profit is $\lambda Q_t P_t - K$. It costs $C$ to close the field, and $J$ to open the field.

Let $\tilde{V}$ denote the value of a closed field, and $\tilde{U}$ the value of an open field. Then:

$$\tilde{U}(p,q,t) = \sup_{\tau}\{E^{p,q,t}[\int_t^{\tau}(\lambda Q_s P_s - K)e^{-rs}ds + \tilde{V}(P_\tau, Q_\tau, \tau) - Ce^{-r\tau}]\} \tag{71}$$

and

$$\tilde{V}(p,q,t) = \sup_{\tau}\{E^{p,q,t}[\tilde{U}(P_\tau, Q_\tau, \tau) - Je^{-r\tau}]\} \tag{72}$$

It is reasonable to guess that:

$$\tilde{V}(p,q,t) = V(p,q)e^{-rt} \tag{73}$$

$$\tilde{U}(p,q,t) = U(p,q)e^{-rt} \tag{74}$$

We search for solution proposals $v$ and $u$ that solve the free boundary problem (59). The first condition (a), states that (using the decomposition (73) and (74)):

$$L_V v = 0 \tag{75}$$

$$L_U u = -\lambda qp + K \tag{76}$$

where

$$L_V = -r + \alpha p\frac{\partial}{\partial p} + \frac{1}{2}\beta^2 p^2 \frac{\partial^2}{\partial p^2} \tag{77}$$

and

$$L_U = L_V - \lambda q \frac{\partial}{\partial q}. \tag{78}$$

It is reasonable to guess that the continuation region for the starting problem is of the form $\{(t,p,q) : p < x(q)\}$, and for the stopping problem $\{(t,p,q) : p > y(q)\}$. This gives a boundary at $p = 0$ for $v$ and at $p = \infty$ for $u$. That $v(0,q) = 0$ is rather obvious. The boundary condition for $u$ is more complicated. It is reasonable to assume that the expected time until it is optimal to close the field approaches $\infty$ as $p$ approaches $\infty$. In the limit the value of the field should be equal to the value of a field with no option to close. We skip the details. Using this argument, we can derive the boundary condition

$$u(p,q) - \left[\frac{pq}{r+\lambda-\alpha} - \frac{K}{r}\right] \to 0 \quad \text{as } p \to \infty \tag{79}$$

Using these boundary conditions on the equations (75) and (76) we can derive the general form of $u$ and $v$ in the continuation areas. (See Bjerkholt and Brekke (1988) [1])

$$u(p,q) = \frac{pq}{r+\lambda-\alpha} - \frac{K}{r} + c_1 \cdot (pq)^\nu \tag{80}$$

$$v(p,q) = c_2(q) \cdot p^\gamma \tag{81}$$

where

$$\gamma = \frac{-(\alpha - \frac{1}{2}\beta^2) + \sqrt{(\alpha - \frac{1}{2}\beta^2)^2 + 2r\beta^2}}{\beta^2} > 1 \tag{82}$$

$$\nu = \frac{-(\alpha - \lambda - \frac{1}{2}\beta^2) - \sqrt{(\alpha - \lambda - \frac{1}{2}\beta^2)^2 + 2r\beta^2}}{\beta^2} < 0. \tag{83}$$

The boundary conditions at the free boundaries $x(q)$ and $y(q)$ are

$$v(x(q),q) = u(x(q),q) - J \tag{84}$$

$$v'_p(x(q),q) = u'_p(x(q),q) \quad \text{"high contact"} \tag{85}$$

$$u(y(q),q) = v(y(q),q) - C \tag{86}$$

$$u'_p(y(q),q) = v'_p(y(q),q) \quad \text{"high contact"} \tag{87}$$

Inserting the general form of $u$ and $v$ we get

$$\frac{x(q)q}{r+\lambda-\alpha} - \frac{K}{r} + c_1 \cdot (x(q)q)^\nu = c_2(q)(x(q))^\gamma + J \tag{88}$$

$$\frac{x(q)q}{r+\lambda-\alpha} + \nu c_1 \cdot (x(q)q)^\nu = \gamma c_2(q)(x(q))^\gamma \tag{89}$$

$$\frac{y(q)q}{r+\lambda-\alpha} - \frac{K}{r} + c_1 \cdot (y(q)q)^\nu = c_2(q)(y(q))^\gamma - C \tag{90}$$

$$\frac{y(q)q}{r+\lambda-\alpha} + \nu c_1 \cdot (y(q)q)^\nu = \gamma c_2(q)(y(q))^\gamma \tag{91}$$

If we guess that the form of the solution is $x(q) = \frac{x}{q}$, $y(q) = \frac{y}{q}$, and $c_2(q) = k_2 \cdot q^\gamma$, the equation system simplifies to :

$$\frac{x}{r+\lambda-\alpha} + c_1 x^\nu = k_2 x^\gamma + (J + \frac{K}{r}) \tag{92}$$

$$\frac{x}{r+\lambda-\alpha} + \nu c_1 x^\nu = \gamma k_2 x^\gamma \tag{93}$$

$$\frac{y}{r+\lambda-\alpha} + c_1 y^\nu = k_2 y^\gamma + (\frac{K}{r} - C) \tag{94}$$

$$\frac{y}{r+\lambda-\alpha} + \nu c_1 y^\nu = \gamma k_2 y^\gamma \tag{95}$$

These are four equations to determine four unknowns, $x$, $y$, $c_1$ and $k_2$. Suppose there exists a solution to these equations, is then the corresponding stopping rule the optimal policy?

[1] The general form of $v$ was first derived in Olsen and Stensland (1988)

Take the case of optimal starting of the field first. We can pick $W = \{x \in \mathbb{R} : x > 0\}$, and the operator is clearly elliptic for $P_t > 0$. Hence it only remains to prove $L_V(u(p,q) - J) \leq 0$ for $p > x(q)$. Using (78) we get

$$\begin{aligned} L_V(u - J) &= rJ + L_U u + \lambda q \frac{\partial u}{\partial q} \\ &= rJ + K - \lambda pq - \lambda[\frac{pq}{r + \lambda - \alpha} + \nu c_1 (pq)^\nu] \\ &= -\frac{r - \alpha}{r + \lambda - \alpha} pq + (rJ + K) + \lambda \nu c_1 (pq)^\nu. \end{aligned} \quad (96)$$

Since $\nu < 0$ it suffices to prove:

$$\frac{r - \alpha}{r + \lambda - \alpha} pq > (rJ + K) \quad (97)$$

and since the left hand side is increasing in $p$ it suffices to prove this inequality for $p = x(q)$. Combining (92) and (93) we find that (97) is equivalent to

$$(r - \alpha\nu) c_1 x^\nu < (r - \alpha\gamma) k_2 x^\gamma \quad (98)$$

which can easily be checked for a specific solution to (92) – (95). Hence, if (98) is satisfied we can conclude that *if $u$ is the optimal value of an open field*, then $v$ is the optimal value of a closed field.

In the case of $u$, we have to prove $L_U(v - C) \leq -f$, and (60). The first inequality is treated by an argument similar to the one above. As for (60) we have that $f \geq -Ke^{-rt}$ hence by the remark following the proof of theorem 2, we conclude that (60) is satisfied.

Hence *if $v$ is the optimal value of a closed field*, then $u$ is the optimal value of the closed field.

To conclude, we have proved that

$$u(p,q)e^{-rt} = \sup_\tau \{E^{p,q,t}[\int_t^\tau (\lambda Q_s P_s - K) e^{-rs} ds + v(P_\tau, Q_\tau) e^{-r\tau} - Ce^{-r\tau}]\} \quad (99)$$

and

$$v(p,q)e^{-rt} = \sup_\tau \{E^{p,q,t}[u(P_\tau, Q_\tau) e^{-r\tau} - Je^{-r\tau}]\} \quad (100)$$

This is similar to the optimality equation in dynamic programming. To complete the proof that $u$ and $v$ are optimal, an optimality equation for sequential optimal stopping is needed. This is a problem for future research.

# References


[1] Adams, R.A. (1975): *Sobolev Spaces*, Academic Press.

[2] Aronson, D.G. (1967): "Bounds for the fundamental solution of a parabolic equation." *Bull AMS*, **73**, pp. 890-96.

[3] Bensoussan, A. and J.L. Lions (1982): *Applications of Variational Inequalities in Stochastic Control*, North-Holland.

[4] Bjerkholt, O. and K.A. Brekke (1988): *Optimal Starting and Stopping Rules for Resource Depletion when Price is Exogenous and Stochastic*, Discussion Paper No. 40, Central Bureau of Statistics, Oslo.

[5] Brekke, K.A. (1989): *Optimal Oil Production and Use of Oil Revenues under Uncertain Oil Prices*, Thesis for the dr. Polit degree, University of Oslo.

[6] Brennan, M.J. and E.S. Schwartz (1985): "Evaluating Natural Resource Investments", *Journal of Business*, **58**, pp. 135-157.

[7] Dixit, A. (1989): "Entry and Exit Decisions under Uncertainty", *Journal of Political Economy*, **97**, pp. 620-638.

[8] Dynkin, E.B. (1965): *Markov Processes*, Vol II, Springer-Verlag.

[9] Friedman, A. (1976): *Stochastic Differential Equations and Applications*, Vol II, Academic Press.

[10] Littman, W., G. Stampacchia and H.F. Weinberger (1963): "Regular points for elliptic equations with discontinuous coefficients", *Ann Scuola Norm. Sup. Pisa*, **17**, pp. 43-77.

[11] Merton, R.C. (1973): "Theory of rational option pricing", *Bell journal of economics and management science.*

[12] McKean, H.P. (1965): "Appendix: A Free Boundary Problem for the Heat Equation Arising From a Problem of Mathematical Economics", *Industrial Management Review*, 6.

[13] Øksendal, B. (1989): *Stochastic Differential Equations*, (Second edition), Springer Verlag.

[14] Øksendal, B. (1990): "The High Contact Principle in Stochastic Control and Stochastic Waves", To appear in *Seminar on Stochastic Processes 1989*, Birkhäuser.

[15] Olsen, T. and G. Stensland (1988) "Optimal Shut Down Decisions in resource Extraction", *Economic Letters*, Forthcoming.

[16] Shiryayev, A.N. (1978): *Optimal Stopping Rules*, Springer-Verlag.

[17] Samuelson, P.A. (1965): "Rational Theory of Warrant Pricing", *Industrial Management Review*, **6**.