Using Fuel to Control a Process to a Goal

by

W.D. Sudderth* and A.P.N. Weerasinghe
University of Minnesota and Iowa State University

Technical Report No. 537

December 1989

*Research supported by National Science Foundation Grant DMS-8801085

Abstract

The problem treated is that of controlling a process

$$X(t) = x_0 + \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s) + A(t)$$

with values in [0,1]. The nonanticipative controls $(\mu(t),\sigma(t))$ are selected from a set $C(x)$ whenever $X(t-) = x$ and the nondecreasing process $A(t)$ is also chosen by the controller and required to satisfy $0 \le A(t) \le y$ where y is a constant representing the initial amount of fuel. The object is to maximize the probability that $X(t)$ reaches 1 and the optimal process is determined when the function

$$\rho(x) = \sup\{\mu/\sigma^2: (\mu,\sigma) \in C(x)\}$$

is monotone and satisfies certain regularity conditions.

AMS 1980 subject classification. Primary 93E20, 60G40.

Key words: Stochastic control, gambling, bold play, timid play, local time.

1. <u>Introduction</u>. A player begins at position $x \in [0,1)$ with initial fuel supply $y \geq 0$. The stochastic processes $X = \{X(t), t \geq 0\}$ available to the player are of the form

$$(1.1) \qquad X(t) = x + \int_0^t \mu(s)ds + \int_0^t \sigma(s)\, dW(s) + A(t)$$

where $\{W(t)\}$ is a standard Brownian motion on some probability space $(\Omega, \underset{\sim}{F}, P)$ adapted to a right-continuous filtration $\{\underset{\sim}{F}_t\} \subset \underset{\sim}{F}$ where each $\underset{\sim}{F}_t$ is independent of $\{W(t+s)-W(t), s \geq 0\}$ and contains all P-null sets. The processes $\mu(t)$ and $\sigma(t)$ are assumed to be real-valued, progressively measurable, and to satisfy

$$(1.2) \qquad \int_0^t (|\mu(s)| + \sigma^2(s))ds < \infty \quad \text{a.s.}$$

for every $t > 0$. The process $A(t)$ is assumed to be nonnegative, nondecreasing, right-continuous, and adapted to $\{\underset{\sim}{F}_t\}$. The number $A(t)$ represents the amount of fuel used up to time t and we require that, for all t, $A(t) \leq y$.

Associated to every $x' \in [0,1]$ is a <u>control set</u> $C(x')$, which is a nonempty subset of $R \times R^+$. The player is required to choose the value of (μ, σ) from $C(x')$ whenever the current position is x'. More precisely, we assume that $(\mu(t), \sigma(t))$ belongs to $C(X(t-))$ for all $t > 0$. (For convenience we have assumed A and, consequently, X to be right-continuous. However, we think of the left-limit $X(t-)$ as being the player's position at times $t > 0$.) Assume finally that the process X is absorbed at 1 whenever $X(t) = 1$ and is absorbed at 0 whenever $X(t) = 0$ and $A(t) = y$.

Let $\Sigma(x,y)$ be the collection of all processes $X = \{X(t)\}$ described above which are available to a player with initial position x and initial fuel supply y and assume that each of these collections is nonempty. The player's objective is to reach 1 and the value function for the problem is defined as

$$V(x,y) = \sup_{X \in \Sigma(x,y)} P[X(t) = 1 \text{ for some } t \geq 0].$$

Notice that the player controls X through the choice of the fuel expenditure

$\{A(t)\}$ as well as the choice of infinitesimal parameters $\{(\mu(t),\sigma(t))\}$.

There are a number of articles on finite fuel problems in which a player controls only the use of fuel. (See, for example, Benes, Shepp, and Witsenhausen (1980), Heath and Sudderth (1974), Karatzas and Shreve (1986).) In these articles it is typically assumed that μ and σ are known constants. Problems are treated in higher dimensions and with different reward structures.

In the special case where the initial fuel y is equal to zero, the fuel process A(t) must also be zero and our problem reduces to a goal problem of the sort studied by Pestien and Sudderth (1985,1988). These authors showed, under regularity conditions, the player should choose (μ,σ) at each x so that μ/σ^2 attains the supremum

$$\rho(x) = \sup\{\mu/\sigma^2 : (\mu,\sigma) \in C(x)\}, \quad 0 \le x \le 1. \tag{1.3}$$

(Here, 0/0 is taken to be $-\infty$.) We conjecture that this choice of (μ,σ) remains optimal for the more general problems considered here and we will verify it under the assumption that ρ is either an increasing or a decreasing function.

Assume for the rest of the paper that the function ρ is continuous and can be written in the form

$$\rho(x) = \mu_0(x)/\sigma_0(x)^2, \quad 0 \le x \le 1, \tag{1.4}$$

where μ_0 and σ_0 are bounded, continuous functions on [0,1], $\inf \sigma_0 > 0$, and $(\mu_0(x),\sigma_0(x)) \in C(x)$ for every $x \in [0,1]$.

There are two extreme strategies for the utilization of fuel which we call <u>bold</u> and <u>timid play</u>, respectively. Bold play refers to the immediate use of all the fuel available or just enough to reach 1, whichever is smaller. (The functions μ_0, σ_0 are used to select infinitesimal parameters for both bold and timid play.) The process X_1 corresponding to bold play at (x,y) is given by

$$X_1(0) = \min(x+y,1), dX_1(t) = \mu_0(X_1(t))dt + \sigma_0(X_1(t))dW(t). \tag{1.5}$$

Timid play refers to the use of fuel only when 1 can be reached immediately or to reflect the X process when it has reached zero and fuel is available. To define the corresponding process, we first introduce the reflected diffusion X^*. This is a process with values in $[0,\infty)$ satisfying

$$(1.6) \qquad X^*(0) = x, \; dX^*(t) = \mu_0(X^*(t))dt + \sigma_0(X^*(t))dW(t) + dL(t),$$

where L is a continuous, nondecreasing process satisfying $L(0) = 0$, L increases only during $\Delta \equiv \{t: X(t) = 0\}$, and the random set Δ has Lebesque measure zero almost surely. In particular,

$$(1.7) \qquad \int_0^\infty 1_{[X^*(s)>0]} dL(s) = 0.$$

The functions μ_0 and σ_0 are from (1.4) and are extended to $[0,\infty)$ by setting $\mu_0(x) = \mu_0(1)$, $\sigma_0(x) = \sigma_0(1)$ for $x > 1$. The existence and uniqueness of weak solutions X^* to (1.6) follow from Theorems 3.1 and 5.3 of Stroock and Varadhan (1971). A detailed proof is given by Athreya and Weerasinghe (1988). Anderson and Orey (1976) have an elegantly simple proof under the assumption that μ_0 and σ_0 are uniformly Lipshitz continuous.

Define the stopping time τ by

$$\tau = \inf\{t \geq 0: X^*(t)+y-L(t) \in \{0\}\cup[1,\infty)\}.$$

Intuitively, τ is the first time that either the player can jump immediately to 1 or has arrived at 0 with no fuel left. The process X_2 corresponding to timid play at (x,y) is now defined by

$$X_2(t) = \begin{cases} X^*(t) & \text{if } t < \tau, \\ 1 & \text{if } t \geq \tau \text{ and } X^*(\tau)+y-L(\tau) \geq 1, \\ 0 & \text{if } t \geq \tau \text{ and } X^*(\tau)+y-L(\tau) = 0. \end{cases}$$

To state our main results, we also introduce the scale function S for the diffusion on [0,1] with infinitesimal parameters μ_0, σ_0. That is,

$$(1.8) \qquad S(x) = \int_0^x \exp(-2\int_0^r \rho(u)du)dr, \quad 0 \le x \le 1.$$

Theorem 1.1. If ρ is nondecreasing, then bold play is optimal and the value function satisfies

$$(1.9) \qquad V(x,y) = \begin{cases} S(x+y)/S(1) & \text{if } x+y < 1, \\ 1 & \text{if } x+y \ge 1. \end{cases}$$

Theorem 1.2. If ρ is nonincreasing, then timid play is optimal and the value function satisfies

$$(1.10) \quad V(x,y) = \begin{cases} 1 - (1-S(x)/S(1-y)\exp\left\{-\int_{1-y}^{1} (1/S(u))du\right\} & \text{if } x+y < 1, \\ 1 & \text{if } x+y \ge 1. \end{cases}$$

If ρ is constant, then it follows from the two theorems that both bold and timid play are optimal as well as many other strategies.

The rest of the paper is mainly devoted to the proofs of the two theorems. The next section presents a verification lemma which is applied in section 3 to prove Theorem 1.1. Section 4 contains the derivation of a hitting distribution for a reflected diffusion and this distribution seems interesting in itself. The result of section 4 and the verification lemma are used in section 5 to prove Theorem 1.2. In the final section, Theorems 1.1 and 1.2 are applied to generalize and sharpen an old result of Heath and Sudderth (1974) on the control of n-dimensional Brownian motion in an annulus.

2. A verification lemma

It is convenient to reformulate the problem in two dimensions with state space $F = \{(x,y): 0 \le x \le 1, 0 \le y < 1\}$, where the x-coordinate represents the player's position and the y-coordinate the fuel supply. (There is no harm in restricting y to be less than 1, because, if $y \ge 1$, the player can jump to 1 immediately from any x.) For each $(x,y) \in F$ and every process $X \in \Sigma(x,y)$ as in

(1.1), define the process Y by

$$Y(t) = y - A(t), \quad t \geq 0. \tag{2.1}$$

Let $\Sigma_2(x,y)$ be the collection of two-dimensional processes (X,Y) obtained in this way.

Here is a simple verification lemma which is adequate for the purposes of this paper. (There are related verification lemmas in [5], [6], [9], [10], and [11]. All of these references with the exception of [6] consider only processes with continuous paths.)

Lemma 2.1. Let Q be a real-valued function defined on an open subset G of R^2 containing F. Assume

(i) Q has continuous second order derivatives,

(ii) $Q(x,y) \geq 0$ if $(x,y) \in F$ and $Q(x,y) \geq 1$ if $x+y \geq 1$,

(iii) for every $(x,y) \in F$ such that $x+y \leq 1$ and every $(\mu,\sigma) \in C(x)$,

(a) $\mu\frac{\partial Q}{\partial x}(x,y) + \frac{1}{2}\sigma^2\frac{\partial^2 Q}{\partial x^2}(x,y) \leq 0$, and

(b) $\frac{\partial Q}{\partial x}(x,y) \leq \frac{\partial Q}{\partial y}(x,y)$.

Then $Q(x,y) \geq V(x,y)$ for all $(x,y) \in F$.

Proof: If $x+y \geq 1$, the conclusion is immediate from (ii) and the definition of V.

So assume $x+y < 1$ and let $(X,Y) \in \Sigma_2(x,y)$ be a process as in (1.1) and (2.1). Because Y is nonincreasing, $Y(t) \leq y \leq x+y < 1$ for all t. Thus the process (X,Y) takes values in the open set $G \cap \{(a,b): b < 1\}$ where Q is smooth by (i). Apply Ito's formula (cf. Meyer [8, III.8.1, p.285]) to get

$$(2.2)\qquad Q(X(t),Y(t)) = Q(X(0),Y(0)) + \int_0^t \frac{\partial Q}{\partial x}(X(s-),Y(s-))\sigma(s)dW(s)$$

$$+ \int_0^t \left(\frac{1}{2}\sigma^2(s)\frac{\partial^2 Q}{\partial x^2} + \mu(s)\frac{\partial Q}{\partial x}\right)(X(s-),Y(s-))ds$$

$$+ \int_0^t \left(\frac{\partial Q}{\partial x} - \frac{\partial Q}{\partial y}\right)(X(s-),Y(s-))dA(s)$$

$$+ \sum_{0<s\le t} [\Delta Q(X(s),Y(s)) - \left(\frac{\partial Q}{\partial x} - \frac{\partial Q}{\partial y}\right)(X(s-),Y(s-))\Delta A(s)]$$

where we have used the following notation and equalities

$$(2.3)\qquad \begin{aligned} &\Delta Q(X(s),Y(s)) = Q(X(s),Y(s))-Q(X(s-),Y(s-)),\\ &\Delta X(s) = X(s)-X(s-) = Y(s-)-Y(s) = -\Delta Y(s) = \Delta A(s). \end{aligned}$$

In (2.2), we have written $Q(X(0),Y(0))$ rather than $Q(x,y)$ because we do not exclude the possibility that fuel is used immediately and

$$X(0) = x + A(0),$$
$$Y(0) = y - A(0).$$

Notice, however, that

$$X(0) + Y(0) = x+y.$$

So, by (iii)(b),

$$(2.4)\qquad Q(x,y) \ge Q(X(0),Y(0)).$$

It likewise follows from (iii)(b) and (2.3) that, for $\tau = \inf\{t:t \ge 0,\ X(t)+Y(t) \ge 1\}$, (2.5) $$\Delta Q(X(s),Y(s)) \le 0, \quad 0 < s \le \tau.$$

Now, because A is monotone and bounded

$$\sum_{0<s\le t} \Delta A(s) < \infty \qquad \text{for } t > 0.$$

Also, because (X,Y) takes values in the compact set $F\cap\{(a,b):b\le y\}$ where $\frac{\partial Q}{\partial x}$ and $\frac{\partial Q}{\partial y}$ are bounded and continuous, it follows that there is a constant C such that

$$|\Delta Q(X(s),Y(s))| \le C\Delta A(s).$$

Thus all the sums in the following identity are finite

(2.6) $$\sum_{0<s\le t} \left[\Delta Q(X(s),Y(s)) - \left(\frac{\partial Q}{\partial x}-\frac{\partial Q}{\partial y}\right)(X(s-),Y(s-)\Delta A(s)\right]$$

$$= \sum_{0<s\le t} [\Delta Q(X(s),Y(s)) - \sum_{0<s\le t} \left(\frac{\partial Q}{\partial x}-\frac{\partial Q}{\partial y}\right)(X(s-),Y(s-))\Delta A(s)],$$

for $0 < t \le \tau$.

Let $\{A^c(t)\}$ be the continuous part of $\{A(t)\}$. Then, for $0 < t \le \tau$,

(2.7) $$\int_0^t \left(\frac{\partial Q}{\partial x}-\frac{\partial Q}{\partial y}\right)(X(s-),Y(s-))dA(s) - \sum_{0<s\le t} \left(\frac{\partial Q}{\partial x}-\frac{\partial Q}{\partial y}\right)(X(s-),Y(s-))\Delta A(s)$$

$$= \int_0^t \left(\frac{\partial Q}{\partial x}-\frac{\partial Q}{\partial y}\right)(X(s-),Y(s-))dA^c(s).$$

Combine (2.2), (2.6), and (2.7) and then use (2.4), (2.5), and (iii) to get

$$Q(X(t\wedge\tau),Y(t\wedge\tau)) = Q(X(0),Y(0)) + \int_0^{t\wedge\tau} \frac{\partial Q}{\partial x}(X(s-),Y(s-))\sigma(s)dW(s)$$

$$+ \int_0^{t\wedge\tau} \left(\frac{1}{2}\,\sigma^2(s)\frac{\partial^2 Q}{\partial x^2} + \mu(s)\frac{\partial Q}{\partial x}\right)(X(s-),Y(s-))ds$$

$$+ \int_0^{t\wedge\tau} \left(\frac{\partial Q}{\partial x}-\frac{\partial Q}{\partial y}\right)(X(s-),Y(s-))dA^c(s)$$

$$+ \sum_{0<s\le t\wedge\tau} \Delta Q(X(s),Y(s))$$

$$\le Q(x,y) + \int_0^{t\wedge\tau} \frac{\partial Q}{\partial x}(X(s-),Y(s-))\sigma(s)dW(s).$$

So, by the optional sampling theorem together with (ii),

$$\begin{aligned} Q(x,y) &\ge EQ(X(t\wedge\tau),Y(t\wedge\tau)) \\ &\ge P[X(s)+Y(s) \ge 1 \text{ for some } s \in [0,t]] \\ &\ge P[X(s) = 1 \text{ for some } s \in [0,t]]. \end{aligned}$$

Let $t\to\infty$ and take the supremum over $(X,Y) \in \Sigma_2(x,y)$ to get $Q(x,y) \ge V(x,y)$. □

3. The proof of Theorem 1.1.

Let $(x,y) \in F$.

If $x+y \ge 1$, then it is obvious that bold play is optimal and $V(x,y) = 1$.

If $x+y < 1$, then the probability of reaching 1 using bold play is given by

$$Q(x,y) = S(x+y)/S(1) \tag{3.1}$$

where S is the scale function of (1.8). This is clear because the bold player jumps immediately to x+y and then employs the diffusion with infinitesimal parameters μ_0, σ_0. Plainly, $Q \le V$ and we need only prove that $Q \ge V$. The proof will be an application of Lemma 2.1.

First notice that the function Q can be extended to a neighborhood of F in such a way that conditions (i) and (ii) of Lemma 2.1 are satisfied. For example, set $\mu_0(x) = \mu_0(0)$, $\sigma_0(x) = \sigma_0(0)$ for $x < 0$ and $\mu_0(x) = \mu_0(1)$, $\sigma_0(x) = \sigma_0(1)$ for $x > 1$. Next define ρ by (1.4) and S by (1.8) for all x. It is easy to check that Q, as defined by (3.1), satisfies (i) and (ii).

It remains to be verified that Q satisfies conditions (iii)(a) and (iii)(b). To that end, let $(x,y) \in F$, $x+y \le 1$, and $(\mu,\sigma) \in C(x)$. Condition (iii)(b) is immediate since, by (3.1),

$$\frac{\partial Q}{\partial x}(x,y) = \frac{\partial Q}{\partial y}(x,y) = S'(x+y)/S(1).$$

For (iii)(a), calculate as follows:

$$\begin{aligned}\mu\frac{\partial Q}{\partial x}(x,y) + \frac{1}{2}\sigma^2\frac{\partial^2 Q}{\partial x^2}(x,y) &= \mu S'(x+y) + \frac{1}{2}\sigma^2 S''(x+y)\\ &\le \sigma^2[\rho(x)S'(x+y) + \tfrac{1}{2}S''(x+y)]\\ &\le \sigma^2[\rho(x+y)S'(x+y) + \tfrac{1}{2}S''(x+y)]\\ &= 0\end{aligned}$$

where the successive lines are, respectively, by (3.1), by (1.3) and the fact that $S'(x+y) \ge 0$, by the same fact together with our assumption that ρ is nondecreasing, and by (1.8).

The proof of Theorem 1.1 is now complete.

4. The hitting distribution for a reflecting diffusion.

For the proof of Theorem 1.2 we need to calculate the probability of winning for timid play. The calculation will not require any assumption that ρ is monotone and will yield a hitting probability of interest in its own right.

The calculation will be carried out with the goal taken to be a fixed positive number a, not necessarily 1. Let $x \geq 0$, $y \geq 0$, and, in this section, write $X(t)$ for the reflecting diffusion $X^*(t)$ defined in (1.6). Define

$$(4.1) \qquad Y(t) = y - L(t), \quad t \geq 0$$

where L is the local time of X at 0 as in (1.6) and (1.7). Introduce the stopping time

$$(4.2) \qquad \tau = \inf\{t \geq 0 : X(t)+Y(t) = a \text{ or } X(t)+Y(t) = 0\}.$$

The processes X and Y have continuous paths. This, together with our assumptions that μ_0 and σ_0 are continuous and $\inf \sigma_0 > 0$, implies that τ is finite almost surely. Notice also that, if $X(\tau)+Y(\tau) = 0$, then $X(\tau) = Y(\tau) = 0$ almost surely because both X and Y are nonnegative.

Here is the probability we wish to compute,

$$(4.3) \qquad Q(x,y) = P[X(\tau)+Y(\tau) = a \mid X(0) = x, \ Y(0) = y],$$

and here is the result.

Theorem 4.1. For $x \geq 0$, $y \geq 0$, and $x+y < a$,

$$(4.4) \quad Q(x,y) = 1 - (1-S(x)/S(a-y))\exp\{-\int_{a-y}^{a} (1/S(v))dv\}.$$

The method of proof is standard. The first step is to show that the function $U(x,y)$, defined as the right side of (4.4), satisfies a certain differential equation. The next step will use Ito's formula.

Lemma 4.2. The function U satisfies

$$\frac{\partial^2 U}{\partial x^2}(x,y) + 2\rho(x)\frac{\partial U}{\partial x}(x,y) = 0 \quad \text{for } x > 0, y > 0, x+y < a,$$

$$(4.5) \qquad \frac{\partial U}{\partial x}(0,y) = \frac{\partial U}{\partial y}(0,y) \qquad \text{for } 0 < y < a,$$

$$U(0,0) = 0,\ U(x,y) = 1 \qquad \text{if } x \ge 0, y \ge 0, x+y = a.$$

Proof: That U satisfies (4.5) can be verified directly. However, we will give a brief account of how (4.5) was solved for those who are interested.

Let U satisfy (4.5) and set $V(x,y) = \frac{\partial U}{\partial x}(x,y)$ so that $\frac{\partial V}{\partial x} = -2\rho(x)V(x,y)$ and, consequently,

$$V(x,y) = A(y)\exp\{-2\int_0^x \rho(r)dr\}$$

where A is an unknown function. Hence,

$$(4.6) \qquad U(x,y) = A(y)\int_0^x \exp\{-2\int_0^r \rho(u)du\}dr + U(0,y).$$

Notice that $\frac{\partial U}{\partial x}(0,y) = V(0,y) = A(y)$ and, therefore,

$$U(0,y) = \int_0^y A(r)dr$$

because $U(0,0) = 0$. Set $F(y) = U(0,y)$ and use the scale function S of (1.7) to rewrite (4.6) as

$$U(x,y) = A(y)S(x) + F(y).$$

Now $U(a-x,x) = 1$ for $0 \le x \le a$ so that F satisfies

$$F(y) + S(a-y)F'(y) = 1, \ 0 < x < a, \ F(0) = 0.$$

This differential equation is easily solved to give

$$F(y) = 1 - \exp\left(-\int_{a-y}^{a} (1/S(u))du\right)$$

and, in particular,

$$A(y) = F'(y) = (1/S(a-y))\exp\left(-\int_{a-y}^{a} (1/S(u))du\right).$$

The formula for U now follows from (4.6). □

To complete the proof of Theorem 4.1, continue to let $U(x,y)$ be the function on the right side of (4.4) and notice that U is smooth on the set $\{(x,y):y<a\}$ because S can be smoothly extended to the whole line. So, by Ito's formula, (1.6), (1.7), and (4.1),

$$(4.7) \quad U(X(t\wedge\tau),Y(t\wedge\tau)) = U(x,y) + \int_0^{t\wedge\tau} \frac{\partial U}{\partial x}(X(s),Y(s))\sigma_0(X(s))dW(s)$$

$$- \int_0^{t\wedge\tau} \left(\frac{\partial U}{\partial x} - \frac{\partial U}{\partial y}\right)(X(s),Y(s))1_{[X(s)=0]}dL(s)$$

$$+ \int_0^{t\wedge\tau} \frac{1}{2}\sigma_0^2(X(s))\left[\frac{\partial^2 U}{\partial x^2} + 2\rho(X(s))\frac{\partial U}{\partial x}\right](X(s),Y(s))ds.$$

By (4.5) the last two integrals in (4.7) vanish. Also, because σ_0 is bounded and $\frac{\partial U}{\partial x}$ is bounded on the set $\{(r,s): 0 \le r \le a, 0 \le s \le y\}$, the first integral is a martingale. Take the expectation in (4.7) and let $t\to\infty$ to get

$$U(x,y) = E[U(X(\tau),Y(\tau)) \mid X(0) = x, \ Y(0) = y].$$

By (4.5), $U(0,0) = 0$ and $U(r,s) = 1$ if $r \geq 0$, $s \geq 0$, $r+s = 1$. Thus $U(x,y)$ is the hitting probability $Q(x,y)$ of (4.3), and the proof of Theorem 4.1 is complete.

We will conclude this section by reformulating the result of Theorem 4.1 in terms of the hitting probability for the "unrestricted part" of a reflecting diffusion. As before, $x \geq 0$ and $X(t)$ is nonnegative and satisfies

$$X(t) = x + \int_0^t \mu_0(X(s))ds + \int_0^t \sigma_0(X(s))dW(s) + L(t) \tag{4.8}$$

where L is continuous, nondecreasing, $L(0) = 0$, L increases only during the set $\Delta = \{t:X(t) = 0\}$, and the random set Δ has Lebesque measure zero. The functions μ_0, σ_0 are continuous on $[0,\infty)$ and $\inf \sigma_0^2 > 0$. As in Anderson and Orey (1976), we define the <u>unrestricted</u> <u>part</u> of the reflecting diffusion to be the process Z given by

$$Z(t) = x + \int_0^t \mu_0(X(s))ds + \int_0^t \sigma_0(X(s))dW(s). \tag{4.9}$$

Thus

$$X(t) = Z(t) + L(t) \tag{4.10}$$

is the Skorohod decomposition of the process X (cf. Chung and Williams [4]).

<u>Theorem</u> 4.2. Let Z be the unrestricted part of a reflecting diffusion X and assume $X(0) = Z(0) = x \geq 0$. Let m, M be constants with $m < x < M$.

(i) If $m \geq 0$, then

$$P[Z(t) \text{ reaches } M \text{ before } m] = \frac{S(x)-S(m)}{S(M)-S(m)}.$$

(ii) If $m < 0$, then

$$P[Z(t) \text{ reaches } M \text{ before } m] = 1 - (1-S(x)/S(M))\exp\{-\int_M^{M-m} (1/S(u))du\}.$$

(Here S is the scale function of (1.8) for $x \geq 0$.)

Proof: Define the stopping time

$$\sigma = \inf\{t \geq 0: Z(t)=m \text{ or } Z(t)=M\}.$$

If $m \geq 0$, then Z and X agree up to time σ and assertion (i) reduces to a well-known formula. If $m < 0$, set $y = -m$ and $a = M+y$. Define the process Y by (4.1) and check that the events [X(t)+Y(t) reaches a before 0] and [Z(t) reaches M before m] are the same. Thus (ii) follows from Theorem 4.1. □

5. The proof of Theorem 1.2. Let $(x,y) \in F$.

If $x+y \geq 1$, then it is obvious that timid play is optimal and $V(x,y) = 1$.

If $x+y < 1$, then the probability of reaching 1 using timid play is, by Theorem 4.1,

$$Q(x,y) = 1 - (1-S(x)/S(1-y))\exp\{-\int_{1-y}^{1} (1/S(u))du\}. \tag{5.1}$$

Clearly, $Q \leq V$ and it suffices to show $Q \geq V$. The proof will be another application of Lemma 2.1.

Recall from section 3 that there is no difficulty in extending ρ continuously and S smoothly so that they are defined for all real x. Thus formula (5.1) can be used to define Q on the open set $G = \{(x,y):y < 1\} \supset F$ and it is easy to check that Q satisfies conditions (i) and (ii) of Lemma 2.1.

Turn now to condition (iii)(a). By Theorem 4.1 and Lemma 4.2,

$$2\rho(x)\frac{\partial Q}{\partial x}(x,y) + \frac{\partial^2 Q}{\partial x^2}(x,y) = 0 \tag{5.2}$$

for $x > 0$, $y > 0$, $x+y < 1$. Moreover, it follows from continuity that (5.2) holds for all $(x,y) \in F$ such that $x+y \leq 1$. Fix such an (x,y) and let $(\mu,\sigma) \in C(x)$. Then

$$\mu\frac{\partial Q}{\partial x}(x,y) + \frac{1}{2}\sigma^2\frac{\partial^2 Q}{\partial x^2}(x,y) \le \sigma^2[\rho(x)\frac{\partial Q}{\partial x}(x,y) + \frac{1}{2}\frac{\partial^2 Q}{\partial x^2}(x,y)]$$

$$= 0,$$

where the inequality is by (1.3) and the fact that $\frac{\partial Q}{\partial x}$ is nonnegative because $S'(x)$ is nonnegative.

The last condition to be checked is (iii)(b). For $(x,y) \in F$ with $x+y \le 1$, let

$$\phi(x,y) = \frac{\partial Q}{\partial x}(x,y) - \frac{\partial Q}{\partial y}(x,y),$$

and we must show $\phi(x,y) \le 0$. A direct calculation from (5.1) yields

$$\text{(5.3)} \quad \phi(x,y) = \frac{\exp\{-\int_{1-y}^{y}(1/S(u))du\}[S(1-y)S'(x)-S(1-y)-S(x)S'(1-y)+S(x)]}{(S(1-y))^2}.$$

Suppose $x > 0$ so that $S(x) > 0$. Then use (5.3) to see that $\phi(x,y) \le 0$ if and only if

$$\frac{S'(x) - 1}{S(x)} \le \frac{S'(1-y)-1}{S(1-y)}.$$

By assumption, $x \le 1-y$ and so it suffices to show that the function

$$g = (S'-1)/S$$

is nondecreasing on $(0,1]$. Now

$$\begin{aligned} g' &= (S''S-(S'-1)S')/S^2 \\ &= (S'[-2\rho S+1-S'])/S^2 \end{aligned}$$

because $S'' + 2\rho S = 0$. Now $S' \ge 0$. So g will be nondecreasing on $(0,1]$ if

$$(5.4) \qquad 2\rho(z)S(z) + S'(z) \le 1 \quad \text{for } 0 \le z \le 1.$$

To check (5.4), first assume that ρ is differentiable and set

$$h = 2\rho S + S'.$$

Then $h(0) = S'(0) = 1$ and

$$h' = 2\rho' S + 2\rho S' + S'' = 2\rho' S \le 0 \text{ on } [0,1]$$

because $S \ge 0$ and, by hypothesis, ρ is nonincreasing.

If ρ is not differentiable, one can, for each $\epsilon > 0$, find ρ_ϵ which is nonincreasing, differentiable, and within ϵ of ρ everywhere on $[0,1]$. By the argument above

$$2\rho_\epsilon(z)S_\epsilon(z) + S_\epsilon'(z) \le 1 \quad \text{for } 0 \le z \le 1,$$

if S_ϵ is defined by (1.7) with ρ_ϵ in place of ρ. Finally, it is easy to see that

$$2\rho_\epsilon S_\epsilon + S_\epsilon' \to 2\rho S + S'$$

as $\epsilon \to 0$.

This completes the proof of Theorem 1.2.

Remark. The proof shows timid play is optimal for a ρ which fails to be nonincreasing so long as inequality (5.4) is satisfied.

6. Application to a problem in R^n.

Consider now the problem of controlling a process moving in an annulus

$$D = \{x \in R^n : r \le |x| \le R\} \text{ where } 0 < r < R \text{ and } |x| = \sqrt{x_1^2 + \ldots + x_n^2}.$$

A player begins at some $x \in D$ with fuel $y \ge 0$. The processes available are of the form

$$(6.1) \qquad X(t) = x + \int_0^t \alpha(s)ds + \int_0^t \beta(s)dW(s) + A(t)$$

where $\{W(t), t \geq 0\}$ is a standard n-dimensional Brownian motion adapted to a filtration $\{\underset{\sim}{F}_t\}$ of complete, right-continuous σ-fields where each $\underset{\sim}{F}_t$ is independent of $\{W(t+s)-W(t), s \geq 0\}$. The functions $\alpha = \alpha(t,w)$ and $\beta = \beta(t,w)$ are assumed to be progressively measurable, adapted to $\{\underset{\sim}{F}_t\}$ with values in R^n and in the space $R^{n\times n}$ of real n×n matrices, respectively, and to satisfy

$$\int_0^t (|\alpha(s)| + |\beta(s)|^2)ds < \infty \text{ a.s. for all } t \geq 0,$$

where $|\beta(s)|^2 = \sum_i \sum_j \beta_{ij}(s)^2$.

It is also required that $(\alpha(t),\beta(t)) \in C(X(t-))$ for every $t > 0$ where the control sets $\{C(x), x \in D\}$ are given subsets of $R^n \times R^{n\times n}$. The process A has values in R^n and is assumed to be right-continuous, adapted to $\{\underset{\sim}{F}_t\}$ and to have paths of total variation at most y. The total variation of A on the interval $[0,t]$ is thought of as the amount of fuel used up to time t. Assume finally that each process is absorbed at X(t) if either $|X(t)| = r$ or $|X(t)| = R$ and the total variation of A on $[0,t]$ is y.

The object of the game is to reach the inner boundary r before arriving at the outer boundary R with no fuel. Let $\Sigma(x,y)$ be the collection of all processes available at (x,y) as described above and assume each such collection is nonempty. The value function for the problem is

$$V(x,y) = \sup_{X\in\Sigma(x,y)} P[|X(t)| = r \text{ for some } t \geq 0].$$

The special case of our problem corresponding to control sets $C(x) = \{(0,I)\}$ so that

$$X(t) = x + W(t) + A(t)$$

was studied by Heath and Sudderth (1974).

The general problem seems too difficult without further assumptions, and we will now assume that the control sets are radially symmetric in the sense that, for each $x \in D$ and each orthogonal $n \times n$ matrix T,

(6.2) $$TC(x) = C(Tx)$$

where TC(x) is defined to be the set $\{(Ta,Tb):(a,b) \in C(x)\}$. Introduce the functions

(6.3) $$\mu(a,b,x) = \frac{\langle x,a\rangle}{|x|} + \frac{\operatorname{trace}(b'b)}{2|x|} - \frac{|b'x|^2}{2|x|^3}$$

$$\sigma(b,x) = \frac{|b'x|}{|x|}$$

where $a \in R^n$, $x \in D$, $b \in R^{n\times n}$, $\langle\cdot,\cdot\rangle$ is the Euclidean inner product, and b' is the transpose of b. Because the Euclidean norm is a smooth function on a neighborhood of D, Ito's formula can be used to see that, for X(t) as in (6.1),

(6.4) $$|X(t)| = |x| + \int_0^t \mu(\alpha(s),\beta(s),X(s-))ds + \int_0^t \sigma(\beta(s),X(s-))dB(s)$$

$$+ \int_0^t \left\langle \frac{X(s-)}{|X(s-)|}, dA(s)\right\rangle + \sum_{0<s\le t} \left[\Delta|X(s)| - \left\langle \frac{X(s-)}{|X(s-)|}, \Delta A(s)\right\rangle\right],$$

where B(t) is the Brownian motion given by

$$B(t) = \int_0^t \left\langle \frac{\beta'(s)X(s)}{|\beta'(s)X(s)|}, dW(s)\right\rangle$$

and $\Delta|X(s)| = |X(s)| - |X(s-)|$, $\Delta A(s) = A(s) - A(s-)$. It is clear from (6.4) that the player is using fuel unwisely unless the process A satisfies

(6.5) $$\left\langle \frac{X(s-)}{|X(s-)|}, dA(s)\right\rangle \le 0, \quad \left\langle \frac{X(s-)}{|X(s-)|}, \Delta A(s)\right\rangle \le 0$$

for all s. This condition, which we now assume, merely says that fuel is used only to move in the desired direction toward the inner boundary.

Define the function η on D by

$$(6.6) \qquad \eta(x) = \inf\{\mu(a,b,x)/\sigma^2(b,x):(a,b) \in C(x)\}.$$

This infimum will play the role here that the supremum of (1.3) played in previous sections. The reason for the change is that our goal is now to move $|X(t)|$ to the left boundary value r in accordance with the formulation in [6]. It follows easily from (6.2) and (6.3) that $\eta(x) = \eta(Tx)$ for T orthogonal. Thus η depends on x only through $|x|$.

We now assume that η can be written in the form

$$(6.7) \qquad \eta(x) = \mu(\alpha_0(x),\beta_0(x),x)/\sigma^2(\beta_0(x),x)$$

where $\alpha_0:D\to R^n$, $\beta_0:D\to R^{n\times n}$ are continuous functions such that $(\alpha_0(x),\beta_0(x)) \in C(x)$ for all $x \in D$ and $\inf \sigma^2(\beta_0(x),x) > 0$. It can be assumed without further loss of generality that $T\alpha_0(x) = \alpha_0(Tx)$, $T\beta_0(x) = \beta_0(Tx)$ for all orthogonal T and all $x \in D$. (If this is not the case already, just fix the values of α_0, β_0 along a certain direction and redefine them elsewhere so as to be orthogonally invariant.) Thus the functions

$$\mu_0(|x|) = \mu(\alpha_0(x),\beta_0(x),x), \quad \sigma_0(|x|) = \sigma(\beta_0(x),x)$$

are well-defined.

As in the one-dimensional case, there are two extreme strategies which we again call <u>bold</u> and <u>timid play</u>, respectively. The functions α_0 and β_0 are used to select the infinitesimal parameters in both cases. Also bold play immediately uses all of the fuel or just enough to reach $\{x: |x| = r\}$ whichever is smaller and, of course, the fuel is used to jump toward the origin. Timid play uses fuel to jump immediately to the goal set $S_r = \{x: |x| = r\}$ when possible and to reflect normally at $S_R = \{x: |x| = R\}$ so long as fuel is available. More precisely, the bold process $X_1 \in \Sigma(x,y)$ satisfies

$$X_1(0) = ((|x|-y)\ r)x/|x|,$$
$$dX_1(t) = \alpha_0(X_1(t))dt + \beta_0(X_1(t))dW(t).$$

To define the timid process $X_2 \in \Sigma(x,y)$, we first define a diffusion process X^* on D with normal reflection on ∂D and satisfying

$$X^*(0) = x$$
$$dX^*(t) = \alpha_0(X^*(t))dt + \beta_0(X^*(t))dW(t) - dK(t)$$

where

$$dK(t) = 1_{[X^*(t)\in\partial D]}n(X^*(t))d|K|(t).$$

Here $K(0) = 0$ and $\{K(t), t\geq 0\}$ is a continuous process adapted to $\{\underset{\sim}{F}_t\}$ and having finite total variation $|K|(t)$ on $[0,t]$ for each t. The vector field $n(\cdot)$ gives the outward unit normal vector field on ∂D. (The existence of such a process X^* follows from results of Stroock and Varadhan (1971). For smooth α_0 and β_0, a simple construction is in Anderson and Orey (1976).)

Apply Ito's formula, as in (6.4), to see that

$$(6.8) \qquad |X^*(t)| = |x| + \int_0^t \mu_0(|X^*(s)|ds) + \int_0^t \sigma_0(|X^*|(s))dB(s)$$
$$- \int_0^t \langle \frac{X^*(s)}{|X^*(s)|}, dK(s)\rangle.$$

Define

$$\tau = \inf\{t\geq 0:\ |X^*(t)| \leq r+y-|K|(t) \text{ or } |X^*(t)| = R \text{ and } |K(t)| = y\}.$$

Notice that, if $0 \leq t < \tau$ and $X^*(t) \in \partial D$, then $|X^*(t)| = R$, $n(X^*(t)) = R^{-1}X^*(t)$ and

$$\langle \frac{X^*(t)}{|X^*(t)|}, dK(t)\rangle = \langle \frac{X^*(t)}{|X^*(t)|}, \frac{X^*(t)}{|X^*(t)|}\rangle d|K(t)| = d|K|(t).$$

Therefore, the local time for the process $|X^*(t)|$ on the boundary $|X^*(t)| = R$ is $|K|(t)$ for $0 \le t \le \tau$, and (6.8) becomes

$$|X^*(t)| = |x| + \int_0^t \mu_0(|X^*(s)|)ds + \int_0^t \sigma_0(|X^*(s)|)dB(s) - |K|(t)$$

for $0 \le t \le \tau$.

Now define the timid process X_2 by

$$X_2(t) = \begin{cases} X^*(t) & \text{if } 0 \le t < \tau, \\ X^*(\tau) & \text{if } t \ge \tau \text{ and } |X^*(\tau)| = R, \\ r\dfrac{X^*(\tau)}{|X^*(\tau)|} & \text{if } t \ge \tau \text{ and } |X^*(\tau)| \le r+y-|K|(\tau). \end{cases}$$

Here is the n-dimensional analogue of Theorems 1.1 and 1.2.

<u>Theorem</u> 6.1. If $\eta(x)$ is nondecreasing (nonincreasing) in $|x|$, then bold play (timid play) is optimal.

<u>Sketch of the proof</u>: Consider a one-dimensional problem on the interval $I = [-R,-r]$ for which the processes available at a point $(z,y) \in I\times[0,\infty)$ correspond to processes $\{-|X(t)|, t \ge 0\}$ where $X \in \Sigma(x,y)$ and $|x| = -z$. The control sets for this problem will be, for $z \in I$, $x \in R^n$, $|x| = -z$,

$$C_1(z) = \{-(\mu(a,b,x),\sigma(b,x)):(a,b) \in C(x)\}.$$

The function ρ defined by

$$\rho(z) = \sup\{\mu/\sigma^2: (\mu,\sigma) \in C_1(z)\}$$

also satisfies $\rho(z) = \eta(x)$ when $|x| = -z$. A change of location and scale now reduces this one-dimensional problem to the problem studied in previous sections. So one need only note that the negative of the radial part of n-dimensional bold (timid) play corresponds to bold (timid) play in one dimension and then apply Theorem 1.1 (1.2). □

It is easy to calculate the value function V from that for the corresponding one-dimensional problem and thereby recover the formula of Heath and Sudderth [6,p.661] in the special case of their problem.

References

[1] Anderson, R.F. and Orey, S. (1976). Small random perturbation of dynamical systems with reflecting boundary. Nagoya Math. Journal 60 189-216.

[2] Athreya, K.B. and Weerasinghe, A.P.N. (1989). Exponentiality of the local time at hitting times for reflecting diffusions and an application. Probability, Statistics, and Mathematics: Papers in Honor of Samuel Karlin. (T.W. Anderson et al, editors) Academic Press, New York.

[3] Benes, V.E., Shepp, L.A., and Witsenhausen, H.S. (1980). Some solvable stochastic control problems. Stochastics 4 39-83.

[4] Chung, K.L. and Williams, R.J. (1983). Introduction to Stochastic Integration. Birkhauser, Boston.

[5] Heath, D., Orey, S., Pestien, V. and Sudderth, W. (1987). Minimizing or maximizing the expected time to reach zero. SIAM J. Control Optim. 25 195-205.

[6] Heath, D.C. and Sudderth, W.D. (1974). Continuous-time gambling problems. Adv. Appl. Prob. 6 651-665.

[7] Karatzas, I. and Shreve, S.E. (1986). Equivalent models for finite fuel stochastic control. Stochastics 18 245-276.

[8] Meyer, P.A. (1974). Un cours sur les integrales stochastiques. Seminaire de Probabilites X. Lecture Notes in Math. 511. Springer-Verlag, Berlin and New York 243-396.

[9] Orey, S., Pestien, V. and Sudderth, W. (1987). Reaching zero rapidly. SIAM J. Control Optim. 25 1253-1265.

[10] Pestien, V. and Sudderth, W. (1985). Continuous-time red and black: how to control a diffusion to a goal. Math. Oper. Res. 13 599-611.

[11] _____ and _____ (1988). Continuous-time casino problems. Math. Oper. Res. 13 364-376.

[12] Stroock, D.W. and Varadhan, S.R.S. (1971). Diffusions with boundary conditions. Comm. Pure. Appl. Math. 24 147-225.