The Annals of Probability
1984, Vol. 12, No. 1, 173–192

GITTINS INDICES IN THE DYNAMIC ALLOCATION PROBLEM FOR DIFFUSION PROCESSES[1]

BY IOANNIS KARATZAS

Columbia University

We discuss the problem of allocating effort among several competing projects, the states of which evolve according to one-dimensional diffusion processes. It is shown that the "play-the-leader" policy of continuing the project with the leading Gittins index is optimal, and very explicit computations of the index are offered. The question of constructing the diffusions according to the above policy is also addressed.

1. Introduction. We consider d "projects" or "investigations", the state of the jth of them at time $t \geq 0$ being denoted by $x_j(t)$; $1 \leq j \leq d$. At each instant of time t, one is allowed to work only on a single project denoted by $i(t)$, which then evolves according to some Markovian rule; meanwhile, the states of all other projects remain frozen. If $i(t) = j$, one acquires an instant reward equal to $h(j, x_j(t))$ per unit time, discounted by the factor $e^{-\alpha t}$. The stochastic control problem is then to choose the "allocation policy" $\{i(t); t \geq 0\}$ in such a way as to maximize the expected discounted reward $E\int_0^\infty e^{-\alpha t}h(i(t), x_{i(t)}(t))\, dt$.

Questions of this sort are often being referred to as "dynamic allocation" problems. They bear a close relationship to a class of statistical questions known as "multi-armed bandit" problems, which have a long history (see the references and the discussion following [8]). The general, discrete-time problem started yielding during the seventies, when Gittins and his collaborators made significant advances in a series of papers (c.f. [7]-[10] and the references therein). Gittins showed that one can associate to each project j an "index" function $M_j(x)$, defined in terms of an optimization problem involving only this particular project and none of the remaining $d - 1$, and such that the following policy is optimal: "at time t, engage the project with the biggest index $M_j(x_j(t))$". The index was first introduced in [10] as the smallest value of a terminal reward which makes immediate stopping profitable when the jth project is in state x, and was later characterized in [9] by means of a "forwards induction" argument. The relatively recent article [15] and book [16] by P. Whittle contain a concise, mathematically rigorous account of the contributions of the Gittins school.

The present work attempts to provide a similar discussion of the dynamic allocation problem in continuous time, when the state of affairs in each one of the d projects can be modeled by a one-dimensional, time-homogeneous diffusion process. Our purpose is to show that the Gittins index can be computed *explicitly*

Received June 1982; revised May 1983.

[1] Research supported by the National Science Foundation under NSF MCS 81-03435.

AMS 1980 *subject classifications.* Primary, 93E20, 60G40; secondary, 60J60.

Key words and phrases. Dynamic allocation, diffusion processes, variational inequalities, optimal stopping, Gittins index, play-the-leader rule.

in this case (c.f. formulae (3.17) and (3.22) in the text), and that a relatively simple argument can be provided for the "forwards induction" characterization (Theorem 4.1). Explicit solutions are also offered for the nonlinear variational inequalities attached to the dynamic allocation problem, which in turn yield the optimality of the index policy as a corollary of Itô's rule.

2. The dynamic allocation problem. Let us assume that, on some probability space $(\Omega, \mathscr{F}, P)$ endowed with the increasing family of σ – fields $\{\mathscr{F}\}_{t\geq o}$, we have a d-dimensional Brownian motion $\{\mathbf{w}(t) = (w_1(t), \cdots, w_d(t)), \mathscr{F}_t; t \geq 0\}$, as well as an "allocation policy" $\mathscr{A} = \{i(t), \mathscr{F}_t; t \geq 0\}$ which is supposed to be a progressively measurable process with values in the set $\{1, \cdots, d\}$. The value of $i(t)$ decides which project is to be engaged at time t, whereas

$$T_j(t) = \text{meas}\{0 \leq s \leq t;\ i(s) = j\}$$

measures the total time-to-date that project j has been in operation. Let us also suppose that, on the same probability space, we have an $\mathbb{R}^d$ – valued process $\mathbf{X} = \{(x_1(t), \cdots, x_d(t)), \mathscr{F}_t; t \geq 0\}$ satisfying the system of stochastic differential equations

$$\begin{aligned} dx_j(t) &= \mu_j(x_j(t))\, dT_j(t) + \sigma_j(x_j(t))\, dw_j(T_j(t)); \quad t \geq 0 \\ x_j(0) &= x_j \end{aligned} \tag{2.1}$$

for $j = 1, \cdots, d$. The process $\{x_j(t); t \geq 0\}$ models the "state of affairs" in project j as a diffusion with local drift $\mu_j(x)$ and variance $\sigma_j^2(x)$ satisfying condition (3.2) below, while the project is being engaged.

REMARK. With a given allocation policy $\mathscr{A}$ and a Brownian Motion $\{\mathbf{B}(t) = (B_1(t), \cdots, B_d(t)), \mathscr{F}_t; t \geq 0\}$ on $(\Omega, \mathscr{F}, P)$, one can construct for each $1 \leq j \leq d$ the solution of the stochastic differential equation with random coefficients

$$dx_j(t) = \mu_j(x_j(t))1_{\{i(t)=j\}}\, dt + \sigma_j(x_j(t))1_{\{i(t)=j\}}\, dB_j(t); \quad t \geq 0 \tag{2.1*}$$

as in [6], page 118. By virtue of a theorem of F. Knight ([11], page 86), the process $\{\mathbf{w}(t), \mathscr{F}_t; t \geq 0\}$ with components

$$w_j(t) = \int_0^{T_j^{-1}(t)} 1_{\{i(s)=j\}}\, dB_j(s); \quad 1 \leq j \leq d$$

is d-dimensional Brownian motion, and so (2.1) is satisfied. □

The reward received by engaging project j is measured by the function $h(j, x)$; $x \in \mathbb{R}$, which we assume to be strictly increasing with bounded, continuous first and second derivatives, and to satisfy

$$\lim_{x\uparrow\infty} h(j, x) = \alpha K, \quad \lim_{x\downarrow-\infty} h(j, x) = \alpha k, \quad \lim_{|x|\uparrow\infty} h'(j, x) = 0 \tag{2.2}$$

for some numbers $K > k$, $\alpha > 0$ and all $1 \leq j \leq d$. Future payoff is discounted by the factor $e^{-\alpha t}$, so that the expected reward corresponding to an initial position $\mathbf{x} = (x_1, \cdots, x_d)$ and an allocation policy $\mathscr{A}$ is given by

$$J(\mathbf{x}; \mathscr{A}) = E_{\mathbf{x}} \int_0^\infty e^{-\alpha t} h(i(t), x_{i(t)}(t))\, dt. \tag{2.3}$$

The *allocation problem* is then to find a policy $\mathscr{A}^*$ so as to maximize expected reward; i.e., with $\Phi(\mathrm{x}) \triangleq \sup_A J(\mathbf{x}; \mathscr{A})$, to achieve $\Phi(\mathbf{x}) = J(\mathbf{x}; \mathscr{A}^*)$, $\forall x \in \mathbb{R}^d$.

This problem can, at least in principle, be treated by methods of dynamic programming. The relevant Bellman equation in $\mathbb{R}^d$:

$$\alpha\hat{\Phi}(\mathbf{x}) = \max_{1\le j\le d}\left[\frac{1}{2}\,\sigma_j^2(x_j)\,\frac{\partial^2}{\partial x_j^2}\,\hat{\Phi}(\mathbf{x}) + \mu_j(x_j)\,\frac{\partial}{\partial x_j}\,\hat{\Phi}(\mathbf{x}) + h(j, x_j)\right], \tag{2.4}$$

is an elliptic, strongly nonlinear partial differential equation of the second order. In order to see its relevance to the allocation problem, let us suppose that a solution $\hat{\Phi}(\mathbf{x})$ of (2.4) is, along with its first derivatives $(\partial/\partial x_j)\hat{\Phi}(\mathbf{x})$; $1 \le j \le d$, bounded and continuous in $\mathbb{R}^d$, with continuous second derivatives $(\partial^2/\partial \mathrm{x}_j^2)\hat{\Phi}(\mathbf{x})$; $1 \le j \le d$. It can be checked by an application of Itô's rule to the process $\{e^{-\alpha t}\hat{\Phi}(\mathbf{x}(t));\ t \ge 0\}$ that $\hat{\Phi}(\mathbf{x})$ is an upper bound on the achievable expected reward:

$$\hat{\Phi}(\mathbf{x}) \ge J(\mathbf{x}; \mathscr{A}), \quad \text{for any allocation policy } \mathscr{A} \text{ and any } \mathbf{x} \in \mathbb{R}^d. \tag{2.5}$$

It is now reasonable to try to construct an allocation policy $\mathscr{A}^*$ in such a way as to achieve equality in relation (2.5); that would in turn entail the optimality of $\mathscr{A}^*$, along with $\Phi(\mathbf{x}) \equiv \hat{\Phi}(\mathbf{x})$; $\forall\, \mathbf{x} \in \mathbb{R}^d$. Evidently, any such attempt hinges on the possibility of solving equation (2.4) explicitly enough to allow the disclosure of the form of the optimal allocation policy from the properties of the solution. In view of the work required to establish even the existence of a classical solution to strongly nonlinear equations (c.f. [3], [4]), and of "the notorious difficulty of obtaining general results in optimal stochastic control theory" as Gittins [7] puts it, it is rather remarkable that such a program is feasible and yields very explicit results.

To make headway, let us consider with P. Whittle [15], [16] the same problem as above, but with the extra option of "retirement", i.e., of abandoning all projects, with an associated fixed payoff M. The expected total reward corresponding to an allocation policy $\mathscr{A}$ and a retirement time τ (stopping time with respect to the family $\{\mathscr{F}_t\}_{t\ge 0}$) becomes

$$J(\mathbf{x}, M; \mathscr{A}, \tau) = E_{\mathbf{x}}\left[\int_0^\tau e^{-\alpha t} h(i(t), x_{i(t)}(t))\, dt + M e^{-\alpha\tau}\right], \tag{2.6}$$

and the allocation problem is to choose $\mathscr{A}^*$, τ^* so that $J(\mathbf{x}, M; \mathscr{A}^*, \tau^*) = F(\mathbf{x}, M)$ for all $\mathbf{x} \in \mathbb{R}^d$ and $M \in \mathbb{R}$, where $F(\mathbf{x}, M) = \sup_{\mathscr{A},\tau} J(\mathbf{x}, M; \mathscr{A}, \tau)$.

As before, in order to reduce the global optimization problem into a pointwise maximization, one has to introduce an appropriate analytical tool, in this case a *nonlinear variational inequality*: to construct, for any fixed M, a function

$\hat{F}(\mathbf{x}, M)$: $\mathbb{R}^d \to \mathbb{R}$ which is bounded and continuous in $\mathbb{R}^d$ along with its gradient, with second partial derivatives $(\partial^2/\partial x_j^2)\hat{F}(\mathbf{x}, M)$ which are continuous off the hyperplanes $S_j \triangleq \{\mathbf{x} \in \mathbb{R}^d;\ x_j = b_j\}$; $1 \le j \le d$, and such that

$$\hat{F}(\mathbf{x}, M) \ge M; \quad \text{in} \quad \mathbb{R}^d, \tag{2.7)(a}$$

$$\max_{1\le j\le d}\left[\frac{1}{2}\sigma_j^2(x_j)\frac{\partial^2}{\partial x_j^2}\hat{F}(\mathbf{x}, M) + \mu_j(x_j)\frac{\partial}{\partial x_j}\hat{F}(\mathbf{x}, M) + h(j, x_j)\right] \le \alpha\hat{F}(\mathbf{x}, M); \quad \text{in} \quad \mathbb{R}^d \setminus \cup_{j=1}^d S_j \tag{2.7)(b}$$

$$[\hat{F}(\mathbf{x}, M) - M]\left[\max_{1\le j\le d}\left\{\frac{1}{2}\sigma_j^2(x_j)\frac{\partial^2}{\partial x_j^2}\hat{F}(\mathbf{x}, M) + \mu_j(x_j)\frac{\partial}{\partial x_j}\hat{F}(\mathbf{x}, M) + h(j, x_j)\right\} - \alpha\hat{F}(\mathbf{x}, M)\right] = 0; \text{ in } \quad \mathbb{R}^d \setminus \cup_{j=1}^d S_j. \tag{2.7)(c}$$

Here, $\mathbf{b} = (b_1, \cdots, b_d)$ is a fixed vector in $\mathbb{R}^d$, with components depending on M.

We shall see in Section 5 that an explicit solution to the variational inequality (2.7) is available in terms of what we call the "Whittle reduction" (relation (5.1)). This device reduces essentially the dynamic allocation problem to d problems of optimal stopping for the component processes viewed independently of one another. It discerns also the form of the optimal allocation policy by assigning an *index function* $M_j(\cdot)$ to the state of each project, and proceeding as follows: "at time t, engage the project with the leading index $M_j(x_j(t))$ as long as the latter exceeds M; otherwise, retire". This reduction also yields the vector $\mathbf{b}$ in the form: $b_j = M_j^{-1}(M)$; $1 \le j \le d$.

In Section 3 we study in detail a generic optimal stopping problem, in terms of which one can introduce the Dynamic Allocation (or Gittins) Index functions $M_j(\cdot)$. Various properties and interpretations of the index are discussed in Section 4. We conclude in Section 6 by showing how to construct the $\mathbb{R}^d$ – valued process $\mathbf{X}$ as in (2.1) on an appropriate probability space, according to our "play-the-leader-index" rule. We discuss this question in the realm of the Stroock-Varadhan theory for diffusion processes; no such problem arises in discrete time.

3. A problem of optimal stopping. Let us consider on a probability space $(\Omega, \mathscr{F}, P; \mathscr{F}_t)$ the one-dimensional diffusion process $X = \{x_t;\ t \ge 0\}$ which satisfies Itô's stochastic differential equation

$$\begin{aligned} dx_t &= \mu(x_t)\,dt + \sigma(x_t)\,dw_t; \quad t \ge 0 \\ x_0 &= x \in \mathbb{R}. \end{aligned} \tag{3.1}$$

The local drift and diffusion coefficients μ and σ are assumed to be $C^1(\mathbb{R})$, with

$$|\sigma'(x)| \le c, \quad \sigma(x) > 0, \quad 0 < \alpha_0 \le \alpha - \mu'(x) \le \beta; \quad \text{all} \quad x \in \mathbb{R}, \tag{3.2}$$

for some positive constants α_0, β and c. The process $W = \{w_t;\ t \geq 0\}$ is Brownian and adapted to the family $\{\mathscr{F}_t\}_{t\geq 0}$. With $h(x)$ a strictly increasing, $C_b^1(\mathbb{R})$ function satisfying conditions (2.2), we consider the following *Optimal Stopping Problem*: to find an $\{\mathscr{F}_t\}$-stopping time τ, so as to maximize the expected discounted reward

$$E_x\left[\int_0^\tau e^{-\alpha t}h(x_t)\,dt + me^{-\alpha\tau}\right].$$

The standard way to attack such problems (see [1] for general theory) is by seeking, for any given $m \in \mathbb{R}$, the $C_b^1(\mathbb{R})$ function $\varphi(x, m)$ with piecewise continuous second derivative $(\partial^2/\partial x^2)\varphi(x, m)$, which solves the variational inequality:

$$\varphi(x, m) \geq m; \quad \text{in } \mathbb{R}, \tag{3.3a}$$

$$\frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2}\varphi(x, m) + \mu(x)\frac{\partial}{\partial x}\varphi(x, m) + h(x) \leq \alpha\varphi(x, m); \quad \text{a.e. in } \mathbb{R}, \tag{3.3b}$$

$$\left[\frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2}\varphi(x, m) + \mu(x)\frac{\partial}{\partial x}\varphi(x, m) + h(x) - \alpha\,\varphi(x, m)\right]\cdot[\varphi(x, m) - m] = 0; \quad \text{a.e. in } \mathbb{R}. \tag{3.3c}$$

It can then be shown by an application of Itô's rule that $\varphi(x, m)$ is the optimal expected reward, and that the optimal stopping time is given by

$$\tau^* = \tau^*(x, m) \triangleq \inf\{t \geq 0;\ \varphi(x_t, m) = m\}. \tag{3.4}$$

We shall seek in this section to solve the variational inequality (3.3) as explicitly as possible.

The increasing nature of the payoff function h suggests that the continuation region for the stopping problem should be an open interval (b, ∞). We look, therefore, for a real constant $b = b(m)$ and a bounded, $C^1(\mathbb{R})$ function $\varphi(x) = \varphi(x, m)$ satisfying the following conditions:

$$\tfrac{1}{2}\sigma^2(x)\varphi''(x) + \mu(x)\varphi'(x) - \alpha\varphi(x) = -h(x); \quad x > b \tag{3.5a}$$

$$< -h(x); \quad x < b \tag{3.5b}$$

$$\varphi(x) = m; \quad x \leq b \quad \text{and} \quad \varphi(x) > m; \quad x > b. \tag{3.5c}$$

Obviously, conditions (3.5) imply (3.3).

Solving equation (3.5a) amounts to obtaining two linearly independent solutions of the corresponding homogeneous equation

$$\tfrac{1}{2}\sigma^2(x)u''(x) + \mu(x)u'(x) - \alpha u(x) = 0, \tag{3.6}$$

as well as a particular solution of the non-homogeneous. For the former, it suffices to recall the pair of functions

$$g(x) = E_x[\exp(-\alpha\tau_0)]; \quad x > 0$$
$$(3.7) \qquad = \frac{1}{E_0[\exp(-\alpha\tau_x)]}; \quad x \le 0$$

and

$$g_1(x) = E_x[\exp(-\alpha\tau_0)]; \quad x \le 0$$
$$(3.8) \qquad = \frac{1}{E_0[\exp(-\alpha\tau_x)]}; \quad x > 0,$$

with the convention $\tau_y = \inf\{t \ge 0; x_t = y\}$; see Itô – McKean [12]. These functions satisfy equation (3.6), are strictly monotone:

$$\lim_{x\uparrow\infty} g(x) = 0, \quad \lim_{x\downarrow-\infty} g(x) = \infty, \quad \lim_{x\uparrow\infty} g_1(x) = \infty, \quad \lim_{x\downarrow-\infty} g_1(x) = 0,$$

and also linearly independent, with Wronskian

$$(3.9) \quad B(x) \triangleq g_1'(x)g(x) - g'(x)g_1(x) = B(0)\exp\left\{-2\int_0^x \frac{\mu(u)}{\sigma^2(u)}\,du\right\} > 0.$$

Besides, for any real numbers $a > b$, we have the composition rules

$$(3.10) \qquad E_a[\exp(-\alpha\tau_b)] = \frac{g(a)}{g(b)}, \quad E_b[\exp(-\alpha\tau_a)] = \frac{g_1(b)}{g_1(a)}.$$

Let us now introduce the diffusion process $Z = \{z_t; t \ge 0\}$ obeying the stochastic differential equation

$$(3.11) \qquad dz_t = [\mu(z_t) + \sigma(z_t)\sigma'(z_t)]\,dt + \sigma(z_t)\,dw_t,$$

and consider the hitting times $T_x = \inf\{t \ge 0; z_t = x\}$ for the process Z. With $z_0 = z > 0$, an application of Itô's rule to the process

$$\left\{g'(z_t)\exp\left(-\int_0^t (a - \mu'(z_s))\,ds\right); 0 \le t \le T_0\right\}$$

yields, in conjunction with equation (3.6) for g,

$$|g'(z)| = |g'(0)|\,E_z\left[\exp\left(-\int_0^{T_0} (\alpha - \mu'(z_s))\,ds\right)\right]$$
$$\le |g'(0)|\,E_z[\exp(-\alpha_0 T_0)] \to_{z\uparrow\infty} 0,$$

while with $z < 0$ we obtain in a similar manner

$$|g'(z)| = \frac{|g'(0)|}{E_0\left[\exp\left(-\int_0^{T_z} (\alpha - \mu'(z_s))\,ds\right)\right]} \ge \frac{|g'(0)|}{E_0[\exp(-\alpha_0 T_z)]} \to_{z\downarrow-\infty} \infty.$$

Accordingly, one can conclude that:

$$\lim_{x\uparrow\infty} g_1'(x) = \infty, \quad \lim_{x\downarrow-\infty} g_1'(x) = 0$$

and

$$\lim_{x\uparrow\infty}\left[g(x)\exp\left\{2\int_0^x \frac{\mu(u)\,du}{\sigma^2(u)}\right\}\right] = 0, \tag{3.12}$$

since

$$g(x)\exp\left\{2\int_0^x \frac{\mu(u)\,du}{\sigma^2(u)}\right\} = \frac{B(0)g(x)}{B(x)} \le \frac{B(0)}{g_1'(x)}.$$

On the other hand, given g one can obtain a linearly independent solution of (3.6) in the form

$$j(x) \triangleq g(x)\int_0^x g^{-2}(y)\exp\left\{-2\int_0^y \frac{\mu(u)\,du}{\sigma^2(u)}\right\} dy;$$

j is of course a positive multiple of g_1 and inherits its growth. It is not hard then to conclude that

$$\lim_{x\uparrow\infty}\left[\,|g'(x)|\exp\left\{2\int_0^x \frac{\mu(u)\,du}{\sigma^2(u)}\right\}\right] = 0. \tag{3.13}$$

A particular solution of the (non-homogeneous) equation (3.5a) is given by the expected reward of continuing forever:

$$p(x) = E_x\int_0^\infty e^{-\alpha t}h(x_t)\,dt.$$

This is again a strictly increasing, $C_b^1(\mathbb{R})$ function, with $k < p(x) < K$ and $\lim_{x\uparrow\infty}p(x) = K$, $\lim_{x\downarrow-\infty}p(x) = k$; its derivative p' admits the stochastic representation

$$\begin{aligned} p'(z) &= E_z\int_0^\infty \exp\left(-\int_0^t (\alpha - \mu'(z_s))\,ds\right)h'(z_t)\,dt \\ &\le E_z\int_0^\infty \exp(-\alpha_0 t)h'(z_t)\,dt, \end{aligned}$$

whence $\lim_{z\to\infty}p'(z) = 0$, by continuity of solutions of stochastic differential equations on their initial conditions (c.f. Friedman [6]).

We can express now the general solution of equation (3.5a) as: $\varphi(x) = Ag(x) + Bg_1(x) + p(x)$. Boundedness of φ on (b, ∞) implies $B = 0$, while elimination of A in $\varphi(b+) = m$ and $\varphi'(b+) = 0$ yields the determining equation for b:

$$M(b) = m, \tag{3.14}$$

where

$$M(x) \triangleq \frac{p'(x)g(x) - g'(x)p(x)}{-g'(x)} \equiv \frac{W(x)}{-g'(x)}.$$

LEMMA 3.1. *Given any m in (k, K), b is uniquely determined by* (3.14).

PROOF. The Wronskian $W(x) = p'(x)g(x) - g'(x)p(x)$ is positive, converges to zero as $x \to \infty$, and besides satisfies the limiting relationship

$$\lim_{x\uparrow\infty}\left[W(x)\exp\left\{2\int_0^x \frac{\mu(u)\,du}{\sigma^2(u)}\right\}\right] = 0 \tag{3.15}$$

(a consequence of (3.12) and (3.13)), as well as the differential equation

$$\tfrac{1}{2}\,\sigma^2(x)W'(x) + \mu(x)W(x) = -h(x)g(x). \tag{3.16}$$

The solution of the latter subject to the condition (3.15) is

$$W(x) = \int_x^\infty \frac{2h(y)g(y)}{\sigma^2(y)}\exp\left\{2\int_x^y \frac{\mu(u)\,du}{\sigma^2(u)}\right\}dy,$$

which yields in turn the following expression for $M(x)$:

$$M(x) = \frac{-2}{g'(x)}\int_x^\infty \frac{h(y)g(y)}{\sigma^2(y)}\exp\left\{2\int_x^y \frac{\mu(u)\,du}{\sigma^2(u)}\right\}dy. \tag{3.17}$$

It is not hard to derive the equation:

$$\frac{\sigma^2(x)}{2}\frac{g'(x)}{g(x)}M'(x) + \alpha M(x) = h(x)$$

and the expression

$$\alpha M(x) = h(x) + \frac{1}{g'(x)}\int_x^\infty h'(y)g'(y)\exp\left\{2\int_x^y \frac{\mu(u)\,du}{\sigma^2(u)}\right\}dy. \tag{3.18}$$

Together, these two yield the formula

$$M'(x) = \frac{-2g(x)}{(\sigma(x)g'(x))^2}\int_x^\infty h'(y)g'(y)\exp\left\{2\int_x^y \frac{\mu(u)\,du}{\sigma^2(u)}\right\}dy \tag{3.19}$$

for the derivative. Clearly then, $M(x)$ is a strictly increasing function, and it can be checked that

$$M(x)\uparrow K\ (M(x)\downarrow k), \quad \text{as} \quad x\uparrow\infty\ (x\downarrow -\infty).$$

The assertion of the lemma is verified.

With b uniquely determined through (3.14) for $k < m < K$, we have

$$\begin{aligned} \varphi(x) &= -\frac{p'(b)}{g'(b)}g(x) + p(x); && x > b \\ &= m; && x \le b. \end{aligned} \tag{3.20}$$

In order to complete the construction of the solution to the stopping problem, we shall verify (3.5b) in the (slightly stronger) form $h(b) < \alpha m$, as well as: $\varphi(x) > m$; $x > b$. The former is a direct consequence of (3.18), since $h(x) < \alpha M(x)$, for all $x \in \mathbb{R}$. For the latter, let us note that $\varphi''(b+) = N(b)/g'(b)$, where $N(x) \triangleq p''(x)g'(x) - p'(x)g''(x)$ satisfies the relation

$$\frac{\sigma^2(x)}{2} N(x) + \alpha W(x) = -h(x)g'(x); \quad x \in \mathbb{R},$$

and observe that (3.18) implies: $\alpha W(x) > -h(x)g'(x)$, for all $x \in \mathbb{R}$. We conclude that $N(x) < 0$ on $\mathbb{R}$, and therefore $\varphi''(b+) > 0$; a fortiori, there exists an $\varepsilon > 0$ so that $\varphi'(x) > 0$ in $(b, b+\varepsilon)$. On the other hand, $\lim_{x\uparrow\infty}\varphi'(x) = 0$; we deduce then that, if φ' were to take a negative value somewhere in (b, ∞), it should also attain a negative minimum in this interval. However, this is impossible by the maximum principle [5], because

$$\begin{aligned} \tfrac{1}{2}\sigma^2(x)(\varphi'(x))'' + (\mu(x) + \sigma(x)\sigma'(x))(\varphi'(x))' - (\alpha - \mu'(x))\varphi'(x) \\ = -h'(x) < 0 \end{aligned}$$

and $\alpha - \mu'(x) \geq \alpha_0 > 0$, for all $x \in \mathbb{R}$. Consequently, $\varphi'(x) \geq 0$, for all $x \geq b + \varepsilon$, whence it follows that $\varphi(x) > m$, all $x > b$.

We have proved the following result:

THEOREM 3.1. *For each $m \in \mathbb{R}$, there exists a unique number $b(m)$ in $\overline{\mathbb{R}}$, so that the optimal stopping time for the problem discussed in this section is given by*

$$\begin{aligned} \tau^* = \tau^*(x, m) &= \inf\{t \geq 0;\ x_t \leq b(m)\} \\ &= +\infty, \quad \textit{if} \quad \{\cdots\} = \varnothing. \end{aligned} \tag{3.21}$$

If $k < m < K$, the number $b(m)$ and the corresponding optimal expected reward $\varphi(x, m)$ are given by (3.14) *and* (3.20), *respectively. If $m \geq K$, then $b(m) = +\infty$, $\tau^* = 0$ and $\varphi(x, m) = m$; for $m \leq k$, we have $b(m) = -\infty$, $\tau^* = +\infty$, and $\varphi(x, m) = p(x) > k$.*

REMARK. In the special case of Brownian motion with drift ($\mu(x) \equiv \mu$, $\sigma(x) \equiv \sigma$), standard computations [12] can show that:

$$g(x) = e^{-\gamma x}, \quad g_1(x) = e^{\beta x}$$

where

$$\beta = \frac{\sqrt{\mu^2 + 2\alpha\sigma^2} - \mu}{\sigma^2}, \quad \gamma = \frac{\sqrt{\mu^2 + 2\alpha\sigma^2} + \mu}{\sigma^2}.$$

Besides,

$$p(x) = E \int_0^\infty e^{-\alpha t} h(x + \mu t + \sigma w_t)\, dt$$
$$= \frac{2}{\sigma^2(\beta + \gamma)} \left[e^{-\gamma x} \int_{-\infty}^x h(y) e^{\gamma y}\, dy + e^{\beta x} \int_x^\infty h(y) e^{-\beta y}\, dy \right]$$

and

$$M(x) = \frac{1}{\alpha} \int_0^\infty h\left(x + \frac{z}{\beta}\right) e^{-z}\, dz. \tag{3.22}$$

4. The Gittins index and its characterizations. The discussion of the properties of the function $M(x)$ in Lemma 3.1 points to the fact that

$$M(x) = \min\{m > k;\ \varphi(x, m) = m\}. \tag{4.1}$$

In other words, $M(x)$ is the smallest value of the terminal reward m which makes immediate stopping profitable, if the diffusion is in state x. This is precisely the Dynamic Allocation Index introduced by Gittins [7], [8], [10] (see also Whittle [15], [16]).

In [9], Gittins and Glazebrook offer yet another interpretation of this index, in terms of a "forwards induction" rule (for a problem in discrete time). We present here a simple derivation of the forwards induction principle for the diffusion case.

LEMMA 4.1. *For all real numbers m and x, we have*

$$\frac{\partial}{\partial m} \varphi(x, m) = E_x[e^{-\alpha \tau^*(x,m)}]. \tag{4.2}$$

PROOF. For $m \le k$, $x \in \mathbb{R}$: $\varphi(x, m) = p(x)$, $\tau^* = \infty$. If $m \ge K$, $x \in \mathbb{R}$, or $k < m < K$, $x \le b(x)$, we have $\varphi(x, m) = m$, $\tau^* = 0$. In either case, there is nothing to be proven. It remains to discuss the case $k < m < K$, $x > b(m)$.

From (3.20) one can check that then

$$\frac{\partial}{\partial m} \varphi(x, m) = -g(x) b'(m) \frac{N(b(m))}{(g'(b(m)))^2}.$$

However, it can be shown that

$$(g'(x))^2 M'(x) = -g(x) N(x)$$

whence, by recalling (3.14), we obtain: $(\partial/\partial m) \cdot \varphi(x, m) = (g(x)/g(b(m)))$. The assertion of the Lemma follows from (3.10). □

THEOREM 4.1. *"Forwards Induction" interpretation of the index. For any $x \in \mathbb{R}$, the Gittins index admits the representation*

$$M(x) = \sup_{\tau>0} \frac{E_x \int_0^{\tau} e^{-\alpha t} h(x_t)\, dt}{1 - E_x e^{-\alpha\tau}}, \tag{4.3}$$

where the supremum is being taken over all $\{\mathscr{F}_t\}$-stopping times τ, such that $P(\tau > 0) = 1$.

PROOF. For any such stopping time, we have from our optimal stopping problem:

$$E_x\left[\int_0^{\tau} e^{-\alpha t} h(x_t)\, dt + M(x)e^{-\alpha\tau}\right] \le \varphi(x, M(x)) = M(x).$$

Therefore,

$$M(x) \ge \frac{E_x \int_0^{\tau} e^{-\alpha t} h(x_t)\, dt}{1 - E_x e^{-\alpha\tau}}$$

for any a.s. positive stopping time τ. In order to obtain a maximizing family of stopping times, let us consider $\tau^*(x, m)$ with $m \uparrow M(x)$. Recalling that for $m < M(x)$ we have

$$\varphi(x, m) = E_x\left[\int_0^{\tau^*(x,m)} e^{-\alpha t} h(x_t)\, dt + me^{-\alpha\tau^*(x,\, m)}\right],$$

we obtain, in conjunction with (4.2):

$$\frac{E_x \int_0^{\tau^*(x,m)} e^{-\alpha t} h(x_t)\, dt}{1 - E_x e^{-\alpha\tau^*(x,m)}} = \frac{\varphi(x, m) - m \dfrac{\partial}{\partial m}\varphi(x, m)}{1 - \dfrac{\partial}{\partial m}\varphi(x, m)}. \tag{4.4}$$

A first attempt to pass to the limit as $m \uparrow M(x)$ leads to an indeterminate form, since $\varphi(x, M(x)) = M(x)$ and $(\partial/\partial m)\varphi(x, M(x)) = 1$. However, let us note

$$\begin{aligned} \frac{\partial^2}{\partial m^2}\varphi(x, m) &= 0; && m \ge M(x) \\ &= \frac{-g'(b(m))g(x)}{g^2(b(m))M'(b(m))}; && m < M(x), \end{aligned} \tag{4.5}$$

so that

$$\frac{\partial^2}{\partial m^2}\varphi(x, M(x)\,-) = \frac{-g'(x)}{g(x)M'(x)} > 0.$$

Using l'Hôpital's rule in (4.4) we thus obtain

$$\lim_{m \uparrow M(x)} \frac{E_x \int_0^{\tau^*(x,m)} e^{-\alpha t} h(x_t)\, dt}{1 - E_x e^{-\alpha \tau^*(x,\, m)}} = M(x). \quad \Box$$

5. The Whittle reduction. We return now to the dynamic allocation problem of Section 2. Let us assume that each pair of drift and diffusion coefficients (μ_j, σ_j); $1 \le j \le d$ satisfies condition (3.2). We consider also the optimal expected reward functions $\varphi_j(x, m)$ and the index functions $M_j(x)$, constructed on the basis of the stopping problem in Section 3 for each diffusion X_j separately, with running payoff function $h(j, \cdot)$, $1 \le j \le d$. Following Whittle [15], we introduce the function

$$\hat{F}(\mathbf{x}, M) \triangleq K - \int_M^K \left(\prod_{i=1}^d \frac{\partial}{\partial m} \varphi_i(x_i, m) \right) dm \tag{5.1}$$

on $\mathbb{R}^d \times \mathbb{R}$; we shall prove that $\hat{F}(\mathbf{x}, M)$ *is a solution of the variational inequality* (2.7).

First, from Lemma 4.1 we have that $0 \le (\partial/\partial m)\varphi_i(x, m) \le 1$; all $x \in \mathbb{R}$, $m \in \mathbb{R}$, $1 \le i \le d$. Consequently,

$$\hat{F}(\mathbf{x}, M) \ge M; \quad \text{for} \quad M \in \mathbb{R}, \quad \mathbf{x} \in \mathbb{R}^d, \tag{5.2}$$

so (a) in (2.7) is satisfied. Secondly, if we denote by $\mathbf{x}^{(i)}$ the vector $(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_d)$ in $\mathbb{R}^{d-1}$ and introduce the functions

$$P_i(\mathbf{x}^{(i)}, m) \triangleq \prod_{j \ne i} \frac{\partial}{\partial m} \varphi_j(x_j, m),$$

we observe (integrating by parts in (5.1) and recalling $\varphi_i(x_i, K) = K$, $P_i(\mathbf{x}^{(i)}, K) = 1$) that

$$\hat{F}(\mathbf{x}, M) = \varphi_i(x_i, M) \cdot P_i(\mathbf{x}^{(i)}, M) + \int_M^K \varphi_i(x_i, m)\, d_m P_i(\mathbf{x}^{(i)}, m),$$

whence

$$\begin{aligned} &\mathscr{L}_i \hat{F}(\mathbf{x}, M) \\ &\quad = \mathscr{L}_i \varphi_i(x_i, M) \cdot P_i(\mathbf{x}^{(i)}, M) + \int_M^K \mathscr{L}_i \varphi_i(x_i, m) \cdot d_m P_i(\mathbf{x}^{(i)}, m), \end{aligned} \tag{5.3}$$

if we denote by $\mathscr{L}_i u$ the differential operator

$$\frac{1}{2} \sigma_i^2(x_i) \frac{\partial^2 u}{\partial x_i^2} + \mu_i(x_i) \frac{\partial u}{\partial x_i} + h(i, x_i) - \alpha u.$$

Recalling (3.5) in the form:

$$\mathscr{L}_i\varphi_i(x_i, M) < 0; \quad M > M_i(x_i)$$
$$= 0; \quad M < M_i(x_i)$$

we obtain in particular

$$(5.4) \qquad \mathscr{L}_i\hat{F}(\mathbf{x}, M) \leq 0; \quad \text{for all } M \in \mathbb{R} \text{ and } \mathbf{x} \in \mathbb{R}^d, \quad \text{such that } M_i(x_i) \neq M,$$

for all $1 \leq i \leq d$, because P_i is non-decreasing in m (see (4.5)). This verifies (b) of (2.7). Finally, in order to check (c), we start by noting that

$$(5.5) \qquad \hat{F}(\mathbf{x}, M) = M; \quad \text{for} \quad M \geq M^*(\mathbf{x}) \triangleq \max_{1\leq j\leq d} M_j(x_j).$$

The proof will be complete if we establish

$$(5.6) \qquad \mathscr{L}_i\hat{F}(\mathbf{x}, M) = 0; \quad \text{for} \quad M < M^*(\mathbf{x}) \equiv M_i(x_i).$$

In this case, $\mathscr{L}_i\varphi_i(x_i, m) = 0$; $M \leq m < M_i(x_i)$ and $P_i(\mathbf{x}^{(i)}, m) = 1$; $M^*(\mathbf{x}^{(i)}) \triangleq \max_{j\neq i} M_j(x_j) \leq m \leq K$. Since $M^*(\mathbf{x}^{(i)}) \leq M^*(\mathbf{x})$, (5.6) follows from (5.3).

Consequently, the function $\hat{F}(\mathbf{x}, M)$ introduced in (5.1) solves the nonlinear variational inequality (2.7). It is similarly checked that the function $\hat{\Phi}(\mathbf{x}) = \hat{F}(\mathbf{x}, k)$ is a classical solution of the Bellman equation (2.4).

Let us now suppose that we can construct, on an appropriate probability space, the allocation rule $\mathscr{A}^* = \{i^*(t), \mathscr{F}_t; t \geq 0\}$ and the d-dimensional diffusion $\mathbf{X}^*$, so that (2.1) is satisfied with

$$(5.7) \qquad i^*(t) = \min\{1 \leq \ell \leq d; M_\ell(x_\ell^*(t)) = \max_{1\leq j\leq d} M_j(x_j^*(t))\}.$$

Then it is not hard to see that this "play the leader" policy of continuing the project with the leading index is indeed optimal for the allocation problem with no retirement option:

$$J(\mathbf{x}; \mathscr{A}^*) = \hat{\Phi}(\mathbf{x}) \equiv \Phi(\mathbf{x}); \quad \forall\ \mathbf{x} \in \mathbb{R}^d.$$

Similarly, one can show that, with $M > k$ and

$$\tau^* = \inf\{t \geq 0; \max_{1\leq j\leq d} M_j(\mathbf{x}_j^*(t)) \leq M\},$$

we have

$$J(\mathbf{x}, M; \mathscr{A}^*, \tau^*) = \hat{F}(\mathbf{x}, M) \equiv F(\mathbf{x}, M); \quad \mathbf{x} \in \mathbb{R}^d.$$

6. Construction of the optimal process. This section is devoted to the construction of the optimal state process $\mathbf{X}$ (we omit the stars) corresponding to the dynamic allocation policy (5.7) of continuing the project with the biggest index. Let us start by partitioning $\mathbb{R}^d$ into the regions

$$Q_1 = \{\mathbf{x} \in \mathbb{R}^d; M_1(x_1) \geq \max_{2\leq i\leq d-1} M_i(x_i)\}$$
$$Q_j = \{\mathbf{x} \in \mathbb{R}^d; M_j(x_j) > \max_{1\leq i\leq j-1} M_i(x_i) \quad \text{and}$$
$$M_j(x_j) \geq \max_{j+1\leq i\leq d} M_i(x_i)\}; \quad 2 \leq j \leq d-1$$
$$Q_d = \{\mathbf{x} \in \mathbb{R}^d; M_d(x_d) > \max_{1\leq i\leq d-1} M_i(x_i)\}.$$

We take as our sample space $\Omega = C([0, \infty); \mathbb{R}^d)$ the space of continuous functions ω on $[0, \infty)$ with values in $\mathbb{R}^d$ and with $\omega(0) = \mathbf{0}$, and define $\mathscr{F}_t = \sigma\{\omega(s); 0 \le s \le t\}$ for every $t \ge 0$. $\mathscr{F} = \sigma\{\omega(s); s \ge 0\}$ is the smallest σ-field which makes all projections measurable, and coincides with the Borel σ-field generated by the topology of uniform convergence on compact subsets. The value of the function ω at time t will be denoted by $\mathbf{x}(t, \omega) = (x_1(t, \omega), \cdots, x_d(t, \omega)) = (x_1(t), \cdots, x_d(t))$. We shall seek to construct a probability measure P on $(\Omega, \mathscr{F})$ and a Brownian motion $\{\mathbf{B}(t) = (B_1(t), \cdots, B_d(t)), \mathscr{F}_t; t \ge 0\}$ on $(\Omega, \mathscr{F}, P)$, so that the system of stochastic equations

$$\begin{aligned} &dx_j(t) \\ (6.1)\qquad &= \mu_j(x_j(t))1_{Q_j}(\mathbf{x}(t))\,dt + \sigma_j(x_j(t))1_{Q_j}(\mathbf{x}(t))\,dB_j(t); \quad t \ge 0,\ 1 \le j \le d \end{aligned}$$

is satisfied, in accordance with (2.1)*. In the Stroock-Varadhan formulation [13], [14], solving (6.1) amounts to constructing a probability measure P on $(\Omega, \mathscr{F})$ such that, with a given point $\mathbf{x} \in \mathbb{R}^d$:

$$(6.2)\ \text{(i)} \qquad P[\mathbf{x}(0) = \mathbf{x}] = 1,$$

and

$$(6.2)\ \text{(ii)} \qquad \left\{f(\mathbf{x}(t)) - \int_0^t Lf(\mathbf{x}(s))\,ds,\ \mathscr{F}_t;\ t \ge 0\right\} \quad \text{is a } P\text{-martingale,}$$

for any real-valued function $f \in C_0^\infty(\mathbb{R}^d)$ (infinitely continuously differentiable with compact support). Here,

$$Lf \triangleq \textstyle\sum_{j=1}^d \left[\mu_j(x_j)\dfrac{\partial f}{\partial x_j} + \dfrac{1}{2}\sigma_j^2(x_j)\dfrac{\partial^2 f}{\partial x_j^2}\right] 1_{Q_j}(\mathbf{x}).$$

This is the so-called *martingale problem* for the diffusion operator L. Because the diffusion matrix for this problem is degenerate, questions of existence and uniqueness to the martingale problem are not immediately covered by the extant theory.

In this section we shall establish the following result:

THEOREM 6.1. *There exists a solution to the martingale problem* (6.2). *Besides, if* $d = 2$, *this solution is unique.*

PROOF OF EXISTENCE. We can remove the degeneracy by considering the sequence of differential operators, indexed by $n \ge 1$:

$$L^{(n)}f \triangleq Lf + \frac{1}{2n}\textstyle\sum_{j=1}^d \dfrac{\partial^2 f}{\partial x_j^2} 1_{Q_j^c}(\mathbf{x}).$$

The new diffusion matrix: $\operatorname{diag}\{a_{jj}^{(n)}(\mathbf{x})\}_{1 \le j \le d}$,

$$a_{jj}^{(n)}(\mathbf{x}) = \sigma_j^2(x_j)1_{Q_j}(\mathbf{x}) + \frac{1}{n} 1_{Q_j^c}(\mathbf{x}),$$

is uniformly positive definite on compact subsets of $\mathbb{R}^d$, with elements admitting—along with the squares of the drift terms—a quadratic growth condition in the space variable. By virtue of Exercise 7.3.2 and Theorem 10.2.2 in [14], there exists for each $n \geq 1$ a probability measure $P^{(n)}$ on $(\Omega, \mathscr{F})$, such that

$$\text{(6.3) (i)} \qquad P^{(n)}[\mathbf{x}(0) = \mathbf{x}] = 1, \quad \text{and}$$

$$\text{(6.3) (ii)} \qquad \left\{ f(\mathbf{x}(t)) - \int_0^t L^{(n)} f(\mathbf{x}(s))\, ds,\ \mathscr{F}_t;\ t \geq 0 \right\} \quad \text{is a } P^{(n)}\text{-martingale,}$$

for all real-valued functions $f \in C_0^\infty(\mathbb{R}^d)$. Now for any $T > 0$, there exists a constant $C(T)$ depending on T and on the various parameters in (3.2), but not on n, such that for all $0 \leq t < t + h \leq T$, $n \geq 1$ we have

$$E^{(n)} |\mathbf{x}(t+h) - \mathbf{x}(t)|^4 \leq C(T)(1 + |\mathbf{x}|^4) h^2;$$

c.f. [6], page 107. By Prokhorov's theorem (Billingsley [2], pages 37 and 95), the sequence $\{P^{(n)}\}_{n=1}^\infty$ is weakly relatively compact. We may assume, therefore, by relabeling indices if necessary, that the sequence $\{P^{(n)}\}_{n=1}^\infty$ converges weakly to a probability measure P on $(\Omega, \mathscr{F})$. From the "portmanteau" theorem of weak convergence ([2], page 12) we see that (6.2) (i) is satisfied.

In order to check (6.2) (ii) it suffices to verify

$$\text{(6.4)} \qquad E\left[\Psi \left\{ f(\mathbf{x}(t)) - f(\mathbf{x}(s)) - \int_s^t Lf(\mathbf{x}(u))\, du \right\}\right] = 0$$

with $0 \leq s < t$, for any bounded, $\mathscr{F}_s$-measurable function $\Psi \colon \Omega \to \mathbb{R}$. It is easily seen, in view of (6.3) (ii), that relation (6.4) is implied by

$$E^{(n)}\left[\Psi \int_s^t L^{(n)} f(\mathbf{x}(u))\, du\right] \to_{n\to\infty} E\left[\Psi \int_s^t Lf(\mathbf{x}(u))\, du\right],$$

or by the convergence to zero as $n \to \infty$ of the sum

$$\text{(6.5)} \qquad \begin{aligned} &E^{(n)}\left[|\Psi| \int_s^t |(L^{(n)} - L) f(\mathbf{x}(u))|\, du\right] \\ &\qquad + \left| E^{(n)}\left[\Psi \int_s^t Lf(\mathbf{x}(u))\, du\right] - E\left[\Psi \int_s^t Lf(\mathbf{x}(u))\, du\right]\right|. \end{aligned}$$

The first term in (6.5) is dominated by const/n, where the constant depends only on the bounds for Ψ and for the second derivatives of f. In order to deal with the second term, we introduce the family of operators $\{L_\delta f;\ \delta > 0\}$ given by

$$L_\delta f \triangleq \sum_{j=1}^d \left[\mu_j(x_j) \frac{\partial f}{\partial x_j} + \frac{1}{2} \sigma_j^2(x_j) \frac{\partial^2 f}{\partial x_j^2}\right] \chi_j^{(\delta)}(\mathbf{x}),$$

where $\chi_j^{(\delta)} \colon \mathbb{R}^d \to [0, 1]$ are continuous functions, such that

$$\begin{aligned} \chi_j^{(\delta)}(\mathbf{x}) &= 1, \quad \text{on} \quad \{\mathbf{x} \in \mathbb{R}^d;\ M_j(x_j) > \delta + \max_{i \neq j} M_i(x_i)\} \subseteq Q_j \\ &= 0, \quad \text{on} \quad Q_j^c \end{aligned}$$

for $j = 1, \cdots, d$. For each $\delta > 0$,

$$\lim_{n\to\infty} E^{(n)}\left[\Psi \int_s^t L_\delta f(\mathbf{x}(u))\,du\right] = E\left[\Psi \int_s^t L_\delta f(\mathbf{x}(u))\,du\right].$$

Therefore, since f has compact support and, for each $j = 1, \cdots, d$:

$$|\chi_j^{(\delta)}(\mathbf{x}) - 1_{Q_j}(\mathbf{x})| \le \sum_{\ell=1,\ell\neq j}^d 1_{\{|M_j(x_j)-M_\ell(x_\ell)|\le\delta\}\cap(Q_j\cup Q_\ell)}(\mathbf{x}),$$

we shall have established (6.4) as soon as we prove that

$$\text{(6.6)}\quad \lim_{\delta\downarrow 0}\sup_{n\ge 1}\max_{1\le j\neq\ell\le d} E^{(n)}\int_0^T 1_{\{|M_j(x_j)-M_\ell(x_\ell)|\le\delta\}\cap\{Q_j\cup Q_\ell\}\cap S}(\mathbf{x}(t))\,dt = 0$$

holds for any compact subset S of $\mathbb{R}^d$ and any $T > 0$.

Towards this end, we introduce the "index processes"

$$\xi_j(t) \triangleq M_j(x_j(t));\quad t \ge 0,\quad 1 \le j \le d$$

which, by the nature of the index functions $M_j(\cdot)$, are bijections of $x_j(t)$; $t \ge 0$ pointwise in time, for each $j = 1, \cdots, d$, and have natural boundaries at the points k, K. We introduce also the function

$$\begin{aligned}\psi(z) &= \tfrac{1}{2} z^2; & 0 \le z \le \delta\\ &= \delta z - \tfrac{1}{2}\delta^2; & z > \delta\\ &= \psi(-z); & z > 0,\end{aligned}$$

for which we have: $\psi(z) \le \delta|z|$, $|\psi'(z)| \le \delta$ and $\psi''(z) = 1_{[-\delta,\delta]}(z)$, for all $z \in \mathbb{R}$. An application of the martingale property (6.3) (ii) to the function $f(\mathbf{x}) = \psi(M_j(x_j) - M_\ell(x_\ell))$, for any fixed integers $j \neq \ell$, $1 \le j, \ell \le d$, gives

$$\begin{aligned}
&E^{(n)}\psi(\xi_j(T) - \xi_\ell(T)) - \psi(\xi_j(0) - \xi_\ell(0))\\
&= E^{(n)}\int_0^T \psi'(\xi_j(t) - \xi_\ell(t))\\
&\qquad \cdot\left[\mu_j(x_j(t))M_j'(x_j(t)) + \frac{1}{2}\sigma_j^2(x_j(t))M_j''(x_j(t))\right]1_{Q_j}(\mathbf{x}(t))\,dt\\
\text{(6.7)}\quad &+ \frac{1}{2n}E^{(n)}\int_0^T \psi'(\xi_j(t) - \xi_\ell(t))M_j''(x_j(t))1_{Q_j^c}(\mathbf{x}(t))\,dt\\
&- E^{(n)}\int_0^T \psi'(\xi_j(t) - \xi_\ell(t))\\
&\qquad \cdot\left[\mu_\ell(x_\ell(t))M_\ell'(x_\ell(t)) + \frac{1}{2}\sigma_\ell^2(x_\ell(t))M_\ell''(x_\ell(t))\right]1_{Q_\ell}Q_\ell(\mathbf{x}(t))\,dt\\
&- \frac{1}{2n}E^{(n)}\int_0^T \psi'(\xi_j(t) - \xi_\ell(t))M_\ell''(x_\ell(t))1_{Q_\ell^c}(\mathbf{x}(t))\,dt
\end{aligned}$$

$$+ \frac{1}{2} E^{(n)} \int_0^T 1_{\{|\xi_j(t)-\xi_\ell(t)|\le\delta\}}$$

$$\cdot \Big[(\sigma_j(x_j(t))M'_j\, x_j(t)))^2 1_{Q_j}(\mathbf{x}(t)) + (\sigma_\ell(x_\ell(t))M'_\ell(x_\ell(t)))^2 1_{Q_\ell}(\mathbf{x}(t))$$

$$+ \frac{1}{n} (M'_j(x_j(t)))^2 1_{Q_j^c}(\mathbf{x}(t)) + \frac{1}{n} (M'_\ell(x_\ell(t)))^2 1_{Q_\ell^c}(\mathbf{x}(t)) \Big] dt$$

for all $n \ge 1$. With S any compact subset of $\mathbb{R}^d$, there exists a positive number $\sigma(S)$ such that

$$\min_{1\le j\le d} \inf_{\mathbf{x}\in S} (\sigma_j(x_j)M'_j(x_j))^2 \ge 2\ \sigma(S) > 0.$$

Applied to the last term on the right-hand side of (6.7), this gives

$$E^{(n)} \int_0^T 1_{\{|M_j(x_j)-M_\ell(x_\ell)|\le\delta\}\cap\{Q_j\cup Q_\ell\}\cap S}(\mathbf{x}(t)) \le \frac{2(K-k)+C(T)}{\sigma(S)}\ \delta$$

for all $n \ge 1$ and $j \ne \ell$, $1 \le j, \ell \le d$, and (6.6) follows.

PROOF OF UNIQUENESS ($d = 2$). It is enough to show that, for any positive constant λ and any real-valued function g which is infinitely continuously differentiable and with compact support in $Q \triangleq (k, K)^2$ off the diagonal $\{(\xi_1, \xi_2); k < \xi_1 = \xi_2 < K\}$, the function

$$u(\xi_1, \xi_2) = E \int_0^\infty e^{-\lambda t} g(\xi_1(t), \xi_2(t))\, dt$$

is uniquely determined, with $\xi_j = \xi_{j(0))} = M_j(x_j(0))$; $j = 1, 2$ (see [14], page 148). To this end, it suffices to show that the resolvent equation

$$\lambda u = \left[b_1(\xi_1) \frac{\partial u}{\partial \xi_1} + \frac{1}{2} s_1^2(\xi_1) \frac{\partial u^2}{\partial \xi_1^2} \right] 1_{\{\xi_1 \ge \xi_2\}}$$

$$(6.8) \qquad + \left[b_2(\xi_2) \frac{\partial u}{\partial \xi_2} + \frac{1}{2} s_2^2(\xi_2) \frac{\partial u^2}{\partial \xi_2^2} \right] 1_{\{\xi_1 < \xi_2\}} + g(\xi_1, \xi_2)$$

admits a solution $u(\xi_1, \xi_2)$ which is continuous on $[k, K]^2$, continuously differentiable in Q and twice continuously differentiable in $\{(\xi_1, \xi_2) \in Q;\ \xi_1 \ne \xi_2\}$, subject to

$$u(K, \xi_2) = 0; \quad k \le \xi_2 \le K \quad \text{and} \quad u(\xi_1, K) = 0; \quad k \le \xi_1 \le K.$$

The functions (b_j, s_j); $j = 1, 2$ are given by

$$b_j(\xi) = M'_j(M_j^{-1}(\xi))\mu_j(M_j^{-1}(\xi)) + ½\, M''_j(M_j^{-1}(\xi))\sigma_j^2(M_j^{-1}(\xi))$$

$$s_j(\xi) = M'_j(M_j^{-1}(\xi))\sigma_j(M_j^{-1}(\xi)).$$

In $Q^+ \triangleq \{(\xi_1, \xi_2) \in Q; \xi_1 > \xi_2\}$ and with $g^+ = g\, 1_{Q^+}$, equation (6.8) becomes

$$\frac{1}{2} s_1^2(\xi_1) \frac{\partial^2 u}{\partial \xi_1^2} + b_1(\xi_1) \frac{\partial u}{\partial \xi_1} - \lambda u = -g^+(\xi_1, \xi_2). \tag{6.8$^+$}$$

If α_+, β_+ are two linearly independent solutions of

$$\tfrac{1}{2} s_1^2(\xi) u''(\xi) + b_1(\xi) u'(\xi) - \lambda u(\xi) = 0$$

and $p^+(\xi_1, \xi_2)$ is a particular solution of $(6.8)^+$, the general solution of the latter in Q^+ is

$$u^+(\xi_1, \xi_2) = c_+(\xi_2)\alpha_+(\xi_1) + d_+(\xi_2)\beta_+(\xi_1) + p^+(\xi_1, \xi_2). \tag{6.9$^+$}$$

Similarly, in $Q^- \triangleq \{(\xi_1, \xi_2) \in Q; \xi_1 < \xi_2\}$ and with $g^- = g\, 1_{Q^-}$,

$$u^-(\xi_1, \xi_2) = c_-(\xi_1)\alpha_-(\xi_2) + d_-(\xi_1)\beta_-(\xi_2) + p^-(\xi_1, \xi_2) \tag{6.9$^-$}$$

is the general solution of equation (6.8) in the form

$$\frac{1}{2} s_2^2(\xi_2) \frac{\partial^2 u}{\partial \xi_2^2} + b_2(\xi_2) \frac{\partial u}{\partial \xi_2} - \lambda u = -\, g^-(\xi_1, \xi_2), \tag{6.8$^-$}$$

where α_-, β_- are two linearly independent solutions of the homogeneous $\tfrac{1}{2} s_2^2(\xi)u''(\xi) + b_2(\xi)u'(\xi) - \lambda u(\xi) = 0$, and $p^-(\xi_1, \xi_2)$ is a particular solution of $(6.8)^-$. The unknown functions $c_\pm$, $d_\pm$ will have to be determined by the continuity and boundary conditions.

By imposing conditions $u^+(K, \xi) = 0$ and $u^-(\xi, K) = 0$ for $k \le \xi \le K$, relations $(6.9)^\pm$ are transformed into

$$u^+(\xi_1, \xi_2) = c_+(\xi_2)\gamma_+(\xi_1) + q^+(\xi_1, \xi_2); \quad \text{in} \quad Q^+ \tag{6.10$^+$}$$

$$u_-(\xi_1, \xi_2) = c_-(\xi_1)\gamma_-(\xi_2) + q^-(\xi_1, \xi_2); \quad \text{in} \quad Q^-, \tag{6.10$^-$}$$

where

$$\gamma_\pm(\xi) \triangleq \alpha_\pm(\xi) - \beta_\pm(\xi) \frac{\alpha_\pm(K)}{\beta_\pm(K)}; \quad k \le \xi \le K,$$

$$q^+(\xi_1, \xi_2) \triangleq p^+(\xi_1, \xi_2) - p^+(K, \xi_2) \frac{\beta_+(\xi_1)}{\beta_+(K)}$$

and

$$q^-(\xi_1, \xi_2) \triangleq p^-(\xi_1, \xi_2) - p^-(\xi_1, K) \frac{\beta_-(\xi_2)}{\beta_-(K)}.$$

Continuity of u and its gradient across the diagonal are tantamount to the conditions

$$c_+(\xi)\gamma_+(\xi) + q^+(\xi, \xi) = c_-(\xi)\gamma_-(\xi) + q^-(\xi, \xi)$$

$$c_+'(\xi)\gamma_+(\xi) + \frac{\partial}{\partial \xi_2} q^+(\xi, \xi) = c_-(\xi)\gamma_-'(\xi) + \frac{\partial}{\partial \xi_2} q^-(\xi, \xi)$$

in $k < \xi < K$, which yield, with $q(\xi_1, \xi_2) \triangleq q^+(\xi_1, \xi_2) - q^-(\xi_1, \xi_2)$, the first-order differential equation

$$c_+'(\xi)\gamma_-(\xi) - c_+(\xi)\gamma_-'(\xi) = \theta(\xi) \triangleq \frac{1}{\gamma_+(\xi)}\left[\gamma_-'(\xi)q(\xi, \xi) - \frac{\partial}{\partial \xi_2} q(\xi, \xi) \cdot \gamma_-(\xi)\right]$$

for $c_+(\cdot)$. The latter is thus given by

$$c_+(\xi) = c_+(k)\,\frac{\gamma_-(\xi)}{\gamma_-(k)} + \gamma_-(\xi)\int_k^\xi \frac{\theta(u)}{\gamma_-^2(u)}\,du; \quad k < \xi < K,$$

and it is readily seen, by reversing the steps of this analysis, that equation (6.8) admits a solution with the desired properties. □

Acknowledgements. The author is indebted to Professor T. L. Lai for introducing him to this topic of research and for continual encouragement and suggestions, to Mr. Fu Chang for a number of stimulating discussions on the multi-armed bandit problem, and to the referee for pointing out two serious flaws in the original typescript. The construction of the optimal process in Section 6 follows detailed suggestions communicated by Professor S. R. S. Varadhan, to whom the author wishes to express his deep appreciation.

REFERENCES

[1] BENSOUSSAN, A. and LIONS, J. L. (1978). *Applications des Inéquations Variationnelles en Contrôle Stochastique.* Dunod, Paris.

[2] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[3] EVANS, L. C. (1982). Classical solutions of fully nonlinear, convex, second-order elliptic equations. *Comm. Pure Appl. Math.* **25** 333–363.

[4] EVANS, L. C. (1983). Classical solutions of the Hamilton-Jacobi-Bellman equation. *Trans. Amer. Math. Soc.* **275** 245–255.

[5] FRIEDMAN, A. (1964). *Partial Differential Equations of Parabolic Type.* Prentice Hall, Englewood Cliffs, N.J.

[6] FRIEDMAN, A. (1975). *Stochastic Differential Equations and Applications, Vol. I.* Academic, New York.

[7] GITTINS, J. C. (1975). The two-armed bandit problem: variations on a conjecture of H. Chernoff. *Sankhyā, Ser. A* **37** 287–291.

[8] GITTINS, J. C. (1979). Bandit processes and dynamic allocation indices (with discussion). *J. Roy. Statist. Soc., Ser. B* **41** 148–177.

[9] GITTINS, J. C. and GLAZEBROOK, K. D. (1977). On Bayesian models in stochastic scheduling. *J. Appl. Probab.* **14** 556–565.

[10] GITTINS, J. C. and JONES, D. M. (1974). A dynamic allocation index in the sequential design of experiments. In *Progress in Statistics* (J. Gani, ed.) 241–266. North Holland, Amsterdam.

[11] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes.* North Holland, Amsterdam.

[12] ITÔ, K. and MCKEAN, H. P. JR. (1974). *Diffusion Processes and their Sample Paths.* Springer, Berlin.

[13] STROOCK, D. W. and VARADHAN, S. R. S. (1969). Diffusion processes with continuous coefficients; Parts I and II. *Comm. Pure Appl. Math.* **22** 345–400 and 479–530.

[14] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes.* Springer, Berlin.

[15] WHITTLE, P. (1980). Multi-armed bandits and the Gittins index. *J. Roy. Statist. Soc., Ser. B.* **42** 143–149.

[16] WHITTLE, P. (1982). *Optimization over Time: Dynamic Programming and Stochastic Control.* Wiley, New York.

DEPARTMENT OF MATHEMATICAL STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, NY 10027