SYNCHRONIZATION AND OPTIMALITY FOR MULTI-ARMED BANDIT PROBLEMS IN CONTINUOUS TIME

by

NICOLE EL KAROUI
Laboratoire de Probabilités
Université Pierre et Marie Curie
4, place Jussieu-Tour 56
75252 Paris, Cedex 05, France

IOANNIS KARATZAS*
Departments of Mathematics and Statistics
619 Mathematics Building
Columbia University
New York, N.Y. 10027, U.S.A.

September 1996

Abstract

We provide a complete solution to a general, continuous-time dynamic allocation (multi-armed bandit) problem with arms that are not necessarily independent or Markovian, using notions and results from time-changes, optimal stopping, and multi-parameter martingale theory. The independence assumption is replaced by the condition (F.4) of Cairoli & Walsh. We also introduce a *synchronization identity* for allocation strategies,

---

* Rsearch supported by the U.S. Army Research Office under Grant DAAH 04-95-I-0528. We are grateful to Prof. Daniel Ocone for finding, and correcting, an error in our original proof of Theorem 7.1.

which is necessary and sufficient for optimality in the case of decreasing rewards, and which leads to the explicit construction of a strategy with all the important properties: optimality in the dynamic allocation problem, optimality in a dual (minimization) problem, and the "index-type" property of Gittins.

Key words and phrases: Dynamic allocation, multi-parameter stochastic calculus and control, optional increasing paths, optimal stopping, time-change, synchronization and "min-plus" algebra.

AMS 1991 subject classifications: Primary 93E20, 60G60, 60H30; secondary 60G40, 90B35, 60G48.

1. INTRODUCTION

Consider a situation, in which several projects are competing simultaneously for the attention of a single investigator (or scarce resource, or equipment of expensive machinery). Let $T_i(t)$ be the total time allocated to the i^{th} project $(i = 1, \ldots, d)$ by the calendar time t, with $\sum_{i=1}^{d} T_i(t) = t$. By engaging the i^{th} project at time t, the investigator accrues a certain reward $h_i(T_i(t))$ per unit time, discounted at the rate $\alpha > 0$ and multiplied by the intensity $\chi_i(t) = \frac{dT_i(t)}{dt}$ with which the project is engaged $(\sum_{i=1}^{d} \chi_i(t) = 1)$. The objective is to schedule the engagement of projects sequentially in time, i.e., to find an *allocation strategy* $\underset{\sim}{T}(t) = (T_1(t), \ldots, T_d(t)),\ t \geq 0$, so as to achieve the maximal total expected discounted reward

$$\Phi \triangleq \sup_{\underset{\sim}{T}(\cdot)} E \sum_{i=1}^{d} \int_0^\infty e^{-\alpha t} h_i(T_i(t)) dT_i(t). \tag{1.1}$$

Here $H_i = \{h_i(u), u \geq 0\}$ is a positive *reward process*, adapted to the *history* $\mathbf{F}_i = \{\mathcal{F}_i(u), u \geq 0\}$ of the corresponding project $i = 1, \ldots, d$. These histories (or *filtrations*) are typically assumed to be independent. A very important example is $h_i(u) = \eta_i(X_i(u))$, where the $X_i(\cdot)$'s are independent Markov processes (modelling the "state of affairs" in each particular project), and the $\eta_i(\cdot)$'s are non-random real-valued functions; we refer to this as the *Markovian case.* One can also envisage situations where the current reward $h_i(u)$ depends on the entire past-and-present record $\mathcal{F}_i(u)$ of the project by time u (e.g., on the best performance-to-date); such situations are clearly *non-Markovian.*

In any case, decisions about engagement of projects can only be made on the basis of available information about the different projects. This is expressed formally by imposing on allocation strategies $\underset{\sim}{T}(\cdot)$ the *non-anticipativity requirement*

$$(1.2) \quad \{T_1(t) \le r_1, \ldots, T_d(t) \le r_d\} \in \mathcal{F}(\underset{\sim}{r}), \quad \forall \; t \ge 0, \quad \forall \; \underset{\sim}{r} = (r_1, \ldots, r_d) \in [0, \infty)^d \,,$$

where $\mathcal{F}(\underset{\sim}{r})$ is the information accumulated on (or, the "collective history" of) all projects $i = 1, \ldots, d$, up to the "multi-dimensional time-parameter" $\underset{\sim}{r} = (r_1, \ldots, r_d) \in [0, \infty)^d$. In the case of projects with indepenent evolution, we take $\mathcal{F}(\underset{\sim}{r}) = \bigvee_{i=1}^d \mathcal{F}_i(r_i)$, where the individual filtrations $\mathbf{F}_i$, $i = 1, \ldots, d$ are assumed to be independent.

This is a generic "Dynamic Allocation" or "Multi-Armed Bandit" problem (the latter terminology is common in the sequential design of experiments). In such problems, projects can be thought of as representing: different arms of a multi-armed bandit machine, that have to be pulled sequentially; medical trials, or chemical research projects, to which effort has to be allocated in order to determine which of d available drugs or treatments is best; different tasks which have to be performed by a single machine or computer; various endeavours or capabilities which have to be deployed sequentially; parcels of agricultural land, which can be cultivated but with limited manpower or resources; and so on (see Gittins (1989)). Models of this sort are designed "to capture the essential conflict inherent in situations, when one has to choose between actions that yield high reward in the short-term, and actions (such as learning, or preparing the ground) whose rewards can be reaped only later," as Whittle (1980) puts it. They are also used in economics, to capture aspects of learning and strategic pricing; see, for example, Rothschild (1974), Banks & Sundaram

(1992), Bergemann & Valimaki (1993).

The problem (1.1) is an apparently difficult question in *stochastic control with multi-dimensional time-parameter*, formulated as in Mandelbaum (1987). The main difficulty comes from the "interaction of the different time-scales" in (1.1). It was shown by Gittins & Jones (1974), Gittins (1979, 1989) and Whittle (1980, 1982) in a discrete-time Markovian context with independently evolving arms, that the computational difficulties of this problem can be reduced to manageable proportions by looking instead at a family of much simpler *optimal stopping problems.* Thus one "splits" the original problem into independent components, and "knits together" the solutions of these latter (as in formula (1.6) below) to obtain the solution of the original problem.

In our general, continuous-time and non-Markovian, setup, and with independent filtrations $\mathbf{F}_i$ $(i = 1, \ldots, d)$, this reduction was carried out in our earlier work El Karoui & Karatzas (1994), as follows: one considers the family of optimal stopping problems

$$e^{-\alpha u} V_i(u;m) \triangleq \operatorname{esssup}_{\substack{\sigma \geq u \\ \sigma \in \mathcal{S}_i}} E\Big[\int_u^\sigma e^{-\alpha\theta} h_i(\theta) d\theta + m e^{-\alpha\sigma} \;\Big|\; \mathcal{F}_i(u)\Big], \quad u \geq 0, \tag{1.3}$$

indexed by the *reward-upon-stopping* $m \geq 0$, where each individual project $i = 1, \ldots, d$ is viewed in isolation (i.e., the supremum in (1.3) is taken over the class $\mathcal{S}_i$ of $\mathbf{F}_i$-stopping times). Let $\sigma_i(u;m) \triangleq \inf\{\theta \geq u / V_i(\theta;m) = m\}$ be the optimal stopping time for this problem: the *Gittins index* ("equitable surrender value") at time u, is then defined as

$$M_i(u) = \inf\{m \geq 0 / V_i(u;m) = m\}\,, \tag{1.4}$$

the smallest value of $m \geq 0$ which makes immediate stopping profitable; and the lower envelope

$$\underline{M}_i(u, \theta) = \inf_{u \leq v \leq \theta} M_i(v), \quad \theta \geq u \tag{1.5}$$

turns out to be the inverse of $m \mapsto \sigma_i(u;m)$. Now we can "knit together" the different optimal rewards in (1.3), to obtain the value of our Dynamic Allocation Problem (1.1) in the form

$$\Phi = \int_0^\infty \Big(1 - \prod_{i=1}^{d} \frac{\partial^+}{\partial m} V_i(0;m)\Big) dm, \tag{1.6}$$

or equivalently

$$\Phi = \int_0^\infty (1 - E e^{-\alpha\tau(m)}) dm = E \int_0^\infty \alpha e^{-\alpha t} N(t) dt. \tag{1.7}$$

Here $\tau(m) \triangleq \sum_{i=1}^{d} \sigma_i(0;m)$, and the inverse $N(t) \triangleq \inf\{m \geq 0 / \tau(m) \leq t\}$ of $m \mapsto \tau(m)$ has an interpretation as "*equitable surrender value for the entire collection of projects*" at calendar time t, in our original problem (1.1). Moreover, an allocation strategy $\underset{\sim}{T}^*(\cdot)$ of *index-type* can be constructed, which "engages projects only from among those with maximal index" so that for every $i = 1, \ldots, d$, and with $\underline{M}_i(\cdot) \equiv \underline{M}_i(0, \cdot)$, the increasing process

$$T_i^*(\cdot) \quad \text{is flat away from the set} \quad \{t \geq 0 / \underline{M}_i(T_i^*(t)) = \max_{1 \leq j \leq d} \underline{M}_j(T_j^*(t))\}. \tag{1.8}$$

This $\underset{\sim}{T}^*(\cdot)$ attains the supremum in (1.1), and is thus optimal in our original problem. This is a very brief synopsis of the results in El Karoui & Karatzas (1994), which builds on and extends the earlier works Karatzas (1984), Mandelbaum (1987), Menaldi & Robin (1990). For similar results in discrete-time, see El Karoui & Karatzas (1993), Cairoli & Dalang (1996), Mandelbaum (1986), Varaiya, Walrand & Buyukkoc (1985), Tsitsiklis (1986, 1994), Kaspi & Mandelbaum (1996), among others.

One of the important points of the present work is to extend these results to situations where the histories $\mathbf{F}_i$ $(i = 1, \ldots, d)$ of the different projects are not necessarily independent. We shall asume instead that the multi-parameter filtration $\mathbf{F} = \{\mathcal{F}(\underset{\sim}{r})\}_{\underset{\sim}{r} \in [0,\infty)^d}$ of (1.2), which expresses "the collective history of the various projects up to the multidimensional time-parameter $\underset{\sim}{r} \in [0, \infty)^d$ ", satisfies the milder condition

$$\mathcal{F}(\underset{\sim}{s}), \ \ \mathcal{F}(\underset{\sim}{r}) \ \text{ are conditionally independent given } \ \mathcal{F}(\underset{\sim}{s} \wedge \underset{\sim}{r}), \quad \forall\ \underset{\sim}{s}, \underset{\sim}{r} \in [0, \infty)^d. \tag{1.9}$$

This condition allows for dependence within the "common past" $\mathcal{F}(\underset{\sim}{s} \wedge \underset{\sim}{r})$ of any two multiparameter time-indices $\underset{\sim}{s}, \underset{\sim}{r} \in [0,\infty)^d$. In the generality of (1.9), the representation (1.6) of the value is no longer true, but it turns out that the representations of (1.7) survive. One can also construct an allocation strategy $\underset{\sim}{T}^*(\cdot)$ of "index-type" (i.e., satisfying (1.8)) which is optimal in this generality. This strategy is shown to satisfy the *"synchronization identity"* (commutation of "minimum" and "summation", or "min-plus" property)

$$\sum_{i=1}^{d} (T_i^*(t) \wedge \sigma_i(0;m)) = t \wedge \tau(m), \quad 0 \le t, m < \infty. \tag{1.10}$$

To our knowledge, this property is being observed and utilised in the context of Dynamic Allocation for the first time in the present work. Indeed, the main feature of this paper is *the central importance of condition (1.10) for Dynamic Allocation*; it is necessary and sufficient for optimality in the special case of decreasing rewards — and leads to all the important results, including the construction of optimal strategies, for the general case as well. It should be of interest, to discover possible connections of (1.10) with the "max-plus" algebra of Baccelli et al. (1992).

The present work owes a lot to the paper of Mandelbaum (1987), and attempts to complete the research programme envisaged and initiated in that paper. For related work on Dynamic Allocation/Multi-Armed Bandit Problems in continuous time, see Dalang (1990), Eplett (1986), Mazziotto & Millet (1987), Kaspi & Mandelbaum (1995), Morimoto (1987), Presman (1990), Presman & Sonin (1983, 1990), Tanaka (1994), Vanderbei (1991), Yushkevich (1988), as well as Chapter 8 of Berry & Fristedt (1985).

2. SUMMARY

The paper is organized as follows. Section 3 presents the complete solution of problem (1.1) in the special but very important case of *decreasing rewards*

$$h_i'(\theta) \stackrel{\triangle}{=} \alpha \underline{M}_i(\theta) \equiv \alpha \underline{M}_i(0,\theta), \quad \theta \ge 0, \quad 1 \le i \le d. \tag{2.1}$$

This solution is *pathwise*, and as such does not impose *any* condition on the multi-parameter filtration $\mathbf{F}$, not even (1.9): for every $\omega \in \Omega$, the supremum of the discounted total reward $\sum_{i=1}^d \int_0^\infty e^{-\alpha t} h_i'(T_i(t,\omega)) dT_i(t,\omega)$ equals

$$\int_0^\infty \alpha e^{-\alpha t} N(t,\omega) dt = \int_0^\infty (1 - e^{-\alpha \tau(m,\omega)}) dm, \tag{2.2}$$

and is attained by any allocation strategy $\underset{\sim}{T}^*(\cdot)$ which satisfies the "synchronization identity" (1.10) (again pathwise). A strategy with this property is constructed explicitly in (3.26) and (3.29); it has the "dual optimality" property

$$\max_{1 \le j \le d} \underline{M}_j(T_j(t)) \ge N(t) = \max_{1 \le i \le d} \underline{M}_j(T_j^*(t)), \quad \forall\; 0 \le t < \infty, \quad \text{for any strategy } \underset{\sim}{T}(\cdot);$$

and is of "index-type", in that it satisfies (1.8) for every $i = 1, \cdots, d$. In particular, the expectation of the random variable in (2.2) gives the value of this problem

(2.3)
$$\Psi \triangleq \sup_{\underset{\sim}{T}(\cdot)} E \sum_{i=1}^d \int_0^\infty e^{-\alpha t} h_i'(T_i(t)) dT_i(t) = E \int_0^\infty \alpha e^{-\alpha t} N(t) dt = E \int_0^\infty (1 - e^{-\alpha \tau(m)}) dm.$$

The results in section 3 are of an "elementary" nature, as they are all based on simple time-change arguments.

The optimality of $\underset{\sim}{T}^*(\cdot)$ in the problem (1.1) with *general reward processes*, is established in section 7. The key steps here are to show that

(i) $\underset{\sim}{T}^*(\cdot)$ has the same expected total discounted reward in both problems (1.1) and (2.3), namely

$$E \sum_{i=1}^d \int_0^\infty e^{-\alpha t} h_i(T_i^*(t)) dT_i^*(t) = E \sum_{i=1}^d \int_0^\infty e^{-\alpha t} \alpha \underline{M_i}(T_i^*(t) dT_i^*(t) = \Psi, \tag{2.4}$$

and that

(ii) the two problems have the same value, namely $\Psi = \Phi$, or equivalently

$$E \sum_{i=1}^d \int_0^\infty e^{-\alpha t} h_i(T_i(t)) dT_i(t) \le \Psi, \quad \text{for any strategy } \underset{\sim}{T}(\cdot). \tag{2.5}$$

The proofs of (2.4), (2.5) are somewhat demanding, as they require

(a) a careful, martingale-type study of the one-parameter *optimal stopping* problem (1.3) (reviewed in section 6), and

(b) notions from the stochastic calculus of *multi-parameter processes*, including filtrations, martingales and optional increasing paths in that context (reviewed in section 4).

In terms of these latter notions, the Dynamic Allocation Problem is formulated in section 5; the critical steps (2.4), (2.5) of its solution are carried out in section 7, *based in a crucial* (perhaps indispensable) *manner on the condition (1.9)*, the famous condition (F.4) of Cairoli & Walsh (1975). Roughly speaking, (1.9) guarantees that (super)martingales of a "small", one-parameter filtration remain (super)martingales of the "large", multi-parameter filtration **F** (see Proposition 4.1 and Lemma 4.6 for precise statements to this effect), and that such multi-parameter (super)martingales retain this property when evaluated along an allocation strategy or optional increasing path (Proposition 4.3).

3. DECREASING REWARDS

The problem (1.1) admits a very general and explicit solution, in the case of *decreasing reward processes* $H_i,\ i = 1, \ldots, d$. In this solution, probability plays no role whatsoever; thus, in order both to simplify typography and to bring out the essentials of this case, let us assume throughout the present section that

$$h_i(u) = \alpha \underline{M}_i(u), \quad 0 < u < \infty \tag{3.1}$$

where $\underline{M}_i : (0,\infty) \to (0,\infty)$ is a given *decreasing, right-continuous function* with $0 < \underline{M}_i(0) \triangleq \lim_{u\downarrow 0} \underline{M}_i(u) \leq \infty$ and $\underline{M}_i(\infty) \triangleq \lim_{u\to\infty} \underline{M}_i(u) = 0$, for every $i = 1, \ldots, d$. An important role will be played by the right-continuous inverses

$$\sigma_i(m) \triangleq \inf\{u \geq 0 / \underline{M}_i(u) \leq m\}, \quad 0 \leq m < \infty \tag{3.2}$$

of these functions, and by the sum of these inverses

$$\tau(m) \triangleq \sum_{i=1}^{d} \sigma_i(m), \quad 0 \leq m < \infty. \tag{3.3}$$

We collect together some elementary properties of all these mappings, for future use.

3.1 Lemma: *For every $i = 1, \ldots, d$ we have the following properties:*

$$\sigma_i(m) \leq u \ \Leftrightarrow \ \underline{M}_i(u) \leq m, \ \forall \ 0 \leq m, u < \infty \tag{3.4}$$

$$\underline{M}_i(\sigma_i(m-)) \leq \underline{M}_i(\sigma_i(m)) \leq m \leq \underline{M}_i(\sigma_i(m)-), \ \ \forall \ 0 < m < \infty \tag{3.5}$$

$$\sigma_i(\underline{M}_i(u-)) \leq \sigma_i(\underline{M}_i(u)) \leq u \leq \sigma_i(\underline{M}_i(u)-), \ \ \forall \ 0 < u < \infty \tag{3.6}$$

$$\mathcal{B}_i \triangleq \{m > 0 / \underline{M}_i(\sigma_i(m-)) < m\} = \bigcup_{t \in \mathbf{B}_i} (\underline{M}_i(t), \underline{M}_i(t-)] \tag{3.7}$$

$$\mathcal{D}_i \triangleq \{u \geq 0 / \sigma_i(\underline{M}_i(u)-) > u\} = \bigcup_{\lambda \in \mathbf{D}_i} [\sigma_i(\lambda), \sigma_i(\lambda-)). \tag{3.8}$$

Here $\mathbf{B}_i$ (respectively, $\mathbf{D}_i$) is the set of discontinuities of the function $\underline{M}_i(\cdot)$ (respectively, $\sigma_i(\cdot)$). The intervals in $\mathcal{B}_i$ (respectively, $\mathcal{D}_i$) are the flat stretches of the function $\sigma_i(\cdot)$ (respectively, $\underline{M}_i(\cdot)$).

3.2 Lemma: *Let us denote the right-continuous inverse of the mapping $m \mapsto \tau(m)$ by*

$$N(u) \triangleq \inf\{m \geq 0 / \tau(m) \leq u\}, \quad 0 \leq u, m < \infty. \tag{3.9}$$

We have then $N(u) > 0, \ \forall \ 0 \leq u < \infty,$ *as well as:*

$$\tau(m) \leq u \ \Leftrightarrow \ N(u) \leq m, \quad \forall \ \ 0 \leq u, m < \infty \tag{3.10}$$

$$(3.11) \qquad N(\tau(m-)) \le N(\tau(m)) \le m \le N(\tau(m)-), \quad \forall \;\; 0 < m < \infty$$

$$(3.12) \qquad \tau(N(u-)) \le \tau(N(u)) \le u \le \tau(N(u)-), \quad \forall \;\; 0 \; < u < \infty$$

$$(3.13) \qquad \mathcal{B} \triangleq \{m > 0 / N(\tau(m-)) < m\} = \bigcup_{t \in \mathbf{B}} (N(t), N(t-)]$$

$$(3.14) \qquad \mathcal{D} \triangleq \{u \ge 0 / \tau(N(u)-) > u\} = \bigcup_{\lambda \in \mathbf{D}} [\tau(\lambda), \tau(\lambda-)).$$

Here $\mathbf{B} = \bigcap_{i=1}^{d} \mathbf{B}_i$ (respectively, $\mathbf{D} = \bigcup_{i=1}^{d} \mathbf{D}_i$) is the set of discontinuities of the function $N(\cdot)$ (respectively, $\tau(\cdot)$). The intervals in $\mathcal{B}$ (respectively, $\mathcal{D}$) are the flat stretches of the function $\tau(\cdot)$ (respectively, $N(\cdot)$).

3.3 Definition: An *allocation strategy* is a vector $\underset{\sim}{T}(\cdot) = (T_1(\cdot), \ldots, T_d(\cdot))$ of increasing functions with $T_1(0) = \ldots = T_d(0) = 0$ and

$$(3.15) \qquad \sum_{i=1}^{d} T_i(t) = t, \quad \forall \; 0 \le t < \infty. \qquad \diamond$$

Every component $T_i(\cdot)$ of an allocation strategy is an absolutely continuous function, since $0 \le T_i(\theta) - T_i(t) \le \sum_{j=1}^{d} (T_j(\theta) - T_j(t)) = \theta - t$ for $0 \le t < \theta < \infty$, $[0, \infty) \to [0, 1]$ with

$$(3.16) \qquad T_i(t) = \int_0^t \chi_i(u) du, \quad 0 \le t < \infty$$

for every $i = 1, \ldots, d$ and

$$(3.17) \qquad \sum_{i=1}^{d} \chi_i(t) = 1, \quad \forall \; 0 \le t < \infty.$$

3.4 Definition: An allocation strategy is called *pure*, if the functions $\chi_i(\cdot), i = 1, \ldots, d$ of (3.16), (3.17) take values in $\{0, 1\}$.

3.5 Interpretation: In the context of Dynamic Allocation, the interpretation is that $T_i(\theta) - T_i(t)$ represents "the total amount of time, during the calendar time-interval $[t, \theta]$, that the allocation strategy $\underset{\sim}{T}(\cdot)$ engages the i^{th} project", and $\chi_i(t)$ represents "the intensity with which $\underset{\sim}{T}(\cdot)$ engages the i^{th} project at the calendar time t". Now $\underset{\sim}{T}(\cdot)$ is pure, if at any time $t \geq 0$ it engages only one project $e(t)$ (i.e., $\chi_{e(t)}(t) = 1$ and $\chi_i(t) = 0, \ \forall \, i \neq e(t)$).

3.6 Dynamic Allocation Problem with Deterministic, Decreasing Rewards. With $h_i(\cdot), i = 1, \ldots, d$ as in (3.1) and given $\alpha \in (0, \infty)$, compute the *value*

$$\Phi \triangleq \sup_{\underset{\sim}{T}(\cdot)} \mathcal{R}(\underset{\sim}{T}), \qquad \text{where} \quad \mathcal{R}(\underset{\sim}{T}) = \sum_{i=1}^{d} \int_0^\infty \alpha e^{-\alpha t} \underline{M}_i(T_i(t)) dT_i(t) \tag{3.18}$$

is the total discounted reward from employing the allocation strategy $\underset{\sim}{T}(\cdot)$, and find an allocation strategy $\underset{\sim}{T}^*(\cdot)$ that attains the supremum (if such a strategy exists).

3.7 Theorem: *The value (3.18) of Problem 3.6 is given by*

$$\Phi = \int_0^\infty \alpha e^{-\alpha u} N(u) du = \int_0^\infty \left(1 - e^{-\alpha \tau(\lambda)}\right) d\lambda; \tag{3.19}$$

and the supremum in (3.18) is achieved by a given allocation strategy $\underset{\sim}{T}(\cdot)$, if and only if $\underset{\sim}{T}(\cdot)$ satisfies the "synchronization identity"

$$\sum_{i=1}^{d} (T_i(t) \wedge \sigma_i(\lambda -)) = t \wedge \tau(\lambda -); \quad \forall \ \ 0 \leq t < \infty, \quad 0 < \lambda < \infty. \tag{3.20}$$

Proof: The results follow from very simple time-change arguments, based on (3.4) and (3.10). First, let us note that $\int_0^\infty \alpha e^{-\alpha u} N(u) du = \int_0^\infty \alpha e^{-\alpha u} \left(\int_0^\infty 1_{\{\lambda < N(u)\}} d\lambda \right) du = \int_0^\infty \left(\int_0^\infty \alpha e^{-\alpha u} 1_{\{u < \tau(\lambda)\}} du \right) d\lambda \Big) = \int_0^\infty (1 - e^{-\alpha \tau(\lambda)}) d\lambda$ gives the second equality in (3.19). Secondly, for any allocation strategy $\underset{\sim}{T}(\cdot)$ let

$$A_i(t; \lambda, \underset{\sim}{T}) \triangleq T_i(t) \wedge \sigma_i(\lambda), \quad A(t; \lambda, \underset{\sim}{T}) \triangleq \sum_{i=1}^{d} A_i(t; \lambda, \underset{\sim}{T}) \leq t \wedge \tau(\lambda) \tag{3.21}$$

for $0 \le t, \lambda < \infty$, and observe:

(3.22)

$$\begin{aligned}
\mathcal{R}(\underset{\sim}{T}) &= \sum_{i=1}^{d} \int_0^\infty \alpha e^{-\alpha t} \underline{M}_i(T_i(t)) dT_i(t) = \sum_{i=1}^{d} \int_0^\infty \alpha e^{-\alpha t} \left(\int_0^\infty 1_{\{\lambda < \underline{M}_i(T_i(t))\}} d\lambda \right) dT_i(t) \\
&= \sum_{i=1}^{d} \int_0^\infty \left(\int_0^\infty \alpha e^{-\alpha t} 1_{\{T_i(t) < \sigma_i(\lambda)\}} dT_i(t) \right) d\lambda \\
&= \sum_{i=1}^{d} \int_0^\infty \int_0^\infty \alpha e^{-\alpha t} dA_i(t; \lambda, \underset{\sim}{T}) d\lambda \\
&= \sum_{i=1}^{d} \int_0^\infty \int_0^\infty \alpha^2 e^{-\alpha t} A_i(t; \lambda, \underset{\sim}{T}) dt d\lambda = \int_0^\infty \int_0^\infty \alpha^2 e^{-\alpha t} A(t; \lambda, \underset{\sim}{T}) dt d\lambda \\
&\le \int_0^\infty \int_0^\infty \alpha^2 e^{-\alpha t} (t \wedge \tau(\lambda)) dt d\lambda = \int_0^\infty \left(1 - e^{-\alpha \tau(\lambda)}\right) d\lambda \ .
\end{aligned}$$

Now (3.22) holds as equality, and thus $\underset{\sim}{T}(\cdot)$ attains the supremum in (3.18), if and only if

$$A(t; \lambda, \underset{\sim}{T}) = t \wedge \tau(\lambda); \quad 0 \le t, \lambda < \infty \tag{3.23}$$

(recall that both sides of (3.23) are continuous in t, and right-continuous in λ). But (3.23) is equivalent to (3.20). $\diamond$

In order to complete the proof of Theorem 3.7, it remains to construct explicitly an allocation strategy that satisfies ther synchronization identity (3.20). To this end, we recall the notation of Lemma 3.2, and start by observing that, for *any allocation strategy* $\underset{\sim}{T}(\cdot)$ *that satisfies (3.20)*, we must have: $\sum_{i=1}^{d}(T_i(\tau(m)) \wedge \sigma_i(m)) = \tau(m)$ $\Big(= \sum_{i=1}^{d} \sigma_i(m) = \sum_{i=1}^{d} T_i(\tau(m))\Big)$, v

$$T_i(\tau(m)) = \sigma_i(m); \quad \forall\ 0 \le m < \infty,\ i = 1, \dots, d. \tag{3.24}$$

On the other hand, for any allocation strategy $\underset{\sim}{T}(\cdot)$ that satisfies (3.24), we have $(T_i \circ \tau)(N(t)) \le T_i(t) \le (T_i \circ \tau)(N(t)-)$ from (3.12) and the increase of $T_i(\cdot)$, and thus

$$\sigma_i(N(t)) \le T_i(t) \le \sigma_i(N(t)-); \quad \forall\ 0 \le t < \infty, \quad i = 1, \dots, d\ . \tag{3.25}$$

3.8 Proposition: *For an allocation strategy $\underset{\sim}{T}(\cdot) = (T_1(\cdot), \ldots, T_d(\cdot))$, the conditions (3.20), (3.24), (3.25) are equivalent.*

Proof: We have already seen that (3.20) $\Rightarrow$ (3.24) $\Rightarrow$ (3.25), so it remains to prove the implication (3.25) $\Rightarrow$ (3.20). With $t \in [0, \infty)$ fixed and $m = N(t)$, we have from (3.25), for $\lambda > m : T_i(t) \geq \sigma_i(m) \geq \sigma_i(\lambda), \ \forall\ i = 1, \ldots, d$. On the other hand, from (3.12) we obtain $t \geq \tau(m) \geq \tau(\lambda)$, whence $\sum_{i=1}^d (T_i(t) \wedge \sigma_i(\lambda)) = \sum_{i=1}^d \sigma_i(\lambda) = \tau(\lambda) = t \wedge \tau(\lambda)$. Similarly, for $\lambda \leq m$, we have from (3.25) that $T_i(t) \leq \sigma_i(m-) \leq \sigma_i(\lambda-), \ \ \forall\ i = 1, \ldots, d,$ and from (3.12) we obtain whence $\sum_{i=1}^d (T_i(t) \wedge \sigma_i(\lambda-)) = \sum_{i=1}^d T_i(t) = t = t \wedge \tau(\lambda-)$. $\diamond$

To proceed further with the construction of a strategy $\underset{\sim}{T}^*(\cdot)$ that satisfies (3.20), we need to distinguish two cases.

Case I: $t \in [0, \infty) \backslash \mathcal{D}$. In this case $\tau(N(t)-) = t$ from (3.14) and (3.12), and writing the identity (3.24) as $T_i^*(\tau(m-)) = \sigma_i(m-)$ with $m = N(t)$, we get

$$T_i^*(t) = \sigma_i(m-) = \sigma_i(N(t)-); \quad i = 1, \ldots, d, \quad t \notin \mathcal{D}. \tag{3.26}$$

Case II: $t \in \mathcal{D}$. Then $t \in L(m) \triangleq [\tau(m), \tau(m-))$ for $m = N(t) \in \mathbf{D}$. We define $y_0 \equiv y_0(m) \triangleq \tau(m)$ and recursively

$$y_i \equiv y_i(m) \triangleq y_{i-1}(m) - \Delta\sigma_i(m) = \sum_{j=1}^{i} \sigma_j(m-) + \sum_{j=i+1}^{d} \sigma_j(m), \quad i = 1, \ldots, d \tag{3.27}$$

where $\Delta\sigma_i(m) \triangleq \sigma_i(m) - \sigma_i(m-)$, and set $L_i(m) \triangleq [y_{i-1}, y_i)$ so that $L(m) = \bigcup_{i=1}^d L_i(m)$. In particular, $y_d = \tau(m-)$, and $L_i(m) = \emptyset$ if $\sigma_i(\cdot)$ is continuous at m. Now find the unique $k = k(t) \in \{1, \ldots, d\}$ for which $t \in L_k(m)$, and write

$$\sum_{i=1}^{d} T_i^*(t) = (t - y_{k-1}) + y_{k-1} = \sum_{i=1}^{k-1} \sigma_i(m-) + (t - y_{k-1} + \sigma_k(m)) + \sum_{i=k+1}^{d} \sigma_i(m) \tag{3.28}$$

(recall (3.25), (3.27)). This suggests taking

$$T_i^*(t) = \left\{ \begin{array}{ccc} \sigma_i(m-) & ; & i = 1, \dots, k-1 \\ \sigma_i(m) + t - y_{i-1} & ; & i = k \\ \sigma_i(m) & ; & i = k+1, \dots, d \end{array} \right\}, \quad \text{for} \quad t \in \mathcal{D}. \tag{3.29}$$

(Similar strategies appear in El Karoui & Karatzas (1994), Proposition 7.3 and in Mandelbaum (1987), p.1537; see also Remark 3.15 below.)

3.9 Proposition: *The vector $\underset{\sim}{T}^*(\cdot) = (T_1^*(\cdot), \dots, T_d^*(\cdot))$ of increasing functions defined in (3.26), (3.29), constitutes an allocation strategy which satisfies the "synchronization" identity (3.20) – and is thus optimal for the Dynamic Allocation Problem (3.18).*

Proof: Each $T_i^*(\cdot)$ is clearly increasing. With $t \notin \mathcal{D}$, we have from (3.26), (3.14): $\sum_{i=1}^d T_i^*(t) = \sum_{i=1}^d \sigma_i(N(t)-) = \tau(N(t)-) = t$. On the other hand, with $t \in \mathcal{D}$, the construction of (3.29) was designed to satisfy (3.28), i.e., $\sum_{i=1}^d T_i^*(t) = t$. Consequently, $\underset{\sim}{T}^*(\cdot)$ is an allocation strategy, and satisfies (3.20) because it satisfies (3.25) rather obviously (cf. Proposition 3.8). $\diamond$

This completes the proof of Theorem 3.7. Let us conclude the section with a few additional properties of strategies $\underset{\sim}{T}^*(\cdot)$ that satisfy (3.20).

3.10 Proposition (Dual Optimality of $\underset{\sim}{T}^*(\cdot)$): *With the notation*

$$\underline{\mathcal{M}}(\underset{\sim}{r}) \stackrel{\triangle}{=} \max_{1 \le i \le d} \underline{M}_i(r_i), \quad \underset{\sim}{r} = (r_1, \dots, r_d) \in [0, \infty)^d \tag{3.30}$$

we have for any strategy $\underset{\sim}{T}(\cdot)$:

$$\underline{\mathcal{M}}(\underset{\sim}{T}(t)) \ge N(t), \quad \forall\ 0 \le t < \infty. \tag{3.31}$$

Equality holds in (3.31) for a given strategy $\underset{\sim}{T}^*(\cdot)$, that is, we have

$$\underline{\mathcal{M}}(\underset{\sim}{T}^*(t)) = N(t), \quad \forall\ 0 \le t < \infty, \tag{3.32}$$

if and only if the strategy $\underset{\sim}{T}^*(\cdot)$ satisfies the synchronization identity (3.20).

Proof: For an arbitrary strategy $\underset{\sim}{T}(\cdot)$, we have the implications

$$\underline{\mathcal{M}}(\underset{\sim}{T}(t)) \leq m \ \Leftrightarrow \ \underline{M}_i(T_i(t)) \leq m, \ \forall\, i = 1, \ldots, d \ \Leftrightarrow \ \sigma_i(m) \leq T_i(t)\,, \ \forall\, i = 1, \ldots, d$$

$$\Rightarrow \ \tau(m) \leq t \ \Leftrightarrow \ N(t) \leq m, \quad \forall \ \ 0 \leq t, m < \infty$$

(consequences of (3.4), (3.10)), and these lead to (3.31). Now (3.31) is valid as equality, if and only if

$$\tau(m) \leq t \quad \Rightarrow \quad \sigma_i(m) \leq T_i(t), \ \ \forall\, i = 1, \ldots, d, \ 0 \leq t, m < \infty. \tag{3.33}$$

But (3.33) holds for any strategy $\underset{\sim}{T}(\cdot)$ that satisfies (3.23) (or, equivalently, (3.20)).

On the other hand, suppose that $\underset{\sim}{T}(\cdot)$ satisfies (3.32); then (3.33) holds as well, and yields: $\sigma_i(m) \leq T_i(\tau(m))$, $\forall\, 0 \leq m < \infty$, $i = 1, \ldots, d$. But $\sum_{i=1}^d \sigma_i(m) = \sum_{i=1}^d T_i(\tau(m)) = \tau(m)$, $\forall \ 0 \leq m < \infty$, so that this $\underset{\sim}{T}(\cdot)$ necessarily satisfies (3.24); from Proposition 3.8, it has to satisfy (3.20) as well. $\diamond$

3.11 Remark: Here is an alternative derivation of (3.32) for a strategy $\underset{\sim}{T}^*(\cdot)$ that satisfies (3.20): from (3.25), (3.5) we have

$$\underline{M}_i(\sigma_i(N(t)-)) \leq \underline{M}_i(T_i^*(t)) \leq \underline{M}_i(\sigma_i(N(t))) \leq N(t), \ \ \forall\ 0 \leq t < \infty, i = 1, \ldots, d \tag{3.34}$$

whence $\max_{i=1,\ldots,d} \underline{M}_i(T_i^*(t)) = \underline{\mathcal{M}}(\underset{\sim}{T}^*(t)) \leq N(t)$; then (3.32) follows from this, in conjunction with (3.31). It develops also that $\underline{M}_i(\sigma_i(N(t))) = N(t)$, for any $i \in \{1, \ldots, d\}$ which attains the maximum.

The following result shows that the support of the increasing function $T_i^*(\cdot)$ is included (modulo Lebesgue measure) in the set $\{t \geq 0 / \underline{M}_i(T_i^*(t)) = \underline{\mathcal{M}}(\underset{\sim}{T}^*(t))\}$, $\ \forall \ \ i = 1, \ldots, d$.

3.12 Proposition: *If a strategy $\underset{\sim}{T}^*(\cdot)$ satisfies the "synchronization identity" (3.20), then*

$$\sum_{i=1}^{d} \int_0^\infty 1_{\{\underline{M}_i(T_i^*(t)) < \underline{\mathcal{M}}(\underset{\sim}{T}^*(t))\}} dT_i^*(t) = 0. \tag{3.35}$$

Proof: From (3.34), (3.7) we have

$$N(t) \notin \mathcal{B}_i \quad \Rightarrow \quad \underline{M}_i(T_i^*(t)) = N(t). \tag{3.36}$$

Consequently, from (3.36) and (3.32), we have for every $i = 1, \ldots, d$:

$$\begin{aligned} 0 &\le \int_0^\infty 1_{\{\underline{M}_i(T_i^*(t)) < \underline{\mathcal{M}}(\underset{\sim}{T}^*(t))\}} dT_i^*(t) = \int_0^\infty 1_{\{\underline{M}_i(T_i^*(t)) < N(t)\}} dT_i^*(t) \\ &\le \int_0^\infty 1_{\{N(t) \in \mathcal{B}_i\}} dT_i^*(t) = 0. \end{aligned}$$

This last equality holds, because $N(t) \in \mathcal{B}_i$ means that $N(t)$ is on a flat stretch of $\sigma_i(\cdot)$, whence that $T_i^*(\cdot) = \sigma_i(N(\cdot)-) = \sigma_i(N(\cdot))$ is then flat at t. $\diamond$

3.13 Remark: *Mixed Dynamic Allocation/Stopping.* The proof of Theorem 3.7 can be slightly modified to show that, with the reward functions of (3.1) and for any given $m \ge 0$, the supremum

$$\Phi(m) \triangleq \sup_{\substack{0 \le \rho \le \infty \\ \underset{\sim}{T}(\cdot)}} \mathcal{R}(\underset{\sim}{T}, \rho; m) \quad \text{with} \quad \mathcal{R}(\underset{\sim}{T}, \rho; m) \triangleq \sum_{i=1}^{d} \int_0^\rho e^{-\alpha t} h_i(T_i(t)) dT_i(t) + m e^{-\alpha \rho}$$

is given as

$$\Phi(m) = \int_0^\infty \alpha e^{-\alpha u} (N(u) \vee m) du = m + \int_m^\infty (1 - e^{-\alpha \tau(\lambda)}) d\lambda,$$

and is attained by any pair of the type by $(\hat{\rho}, \hat{\underset{\sim}{T}}(\cdot))$, where $\hat{\rho} = \tau(m)$ and $\hat{\underset{\sim}{T}}(\cdot)$ a strategy which satisfies (3.20) for all $\lambda \in (m, \infty)$. Note that for any such $\hat{\underset{\sim}{T}}(\cdot)$, the recipe

$$\hat{T}_i(t; m) \triangleq \hat{T}_i(t) \wedge \sigma_i(m), \; i = 1, \ldots, d \quad \text{for} \quad 0 \le t \le \tau(m)$$

defines an allocation strategy of "write-off" type, as in Whittle (1980).

3.14 Remark: For every given allocation strategy $\underset{\sim}{T}(\cdot)$, there exists a sequence *of pure* strategies $\{\underset{\sim}{T}^{(n)}(\cdot)\}_{n\in\mathbf{N}}$ (Definition 3.4) such that $\lim_{n\to\infty}\sup_{0\le t\le T}||\underset{\sim}{T}^{(n)}(t)-\underset{\sim}{T}(t)||=0$. On the other hand, we have from (3.21) and (3.22):

$$\mathcal{R}(\underset{\sim}{T})=\sum_{i=1}^{d}\int_0^\infty\left(\int_0^\infty \alpha^2 e^{-\alpha t}(T_i(t)\wedge\sigma_i(\lambda))dt\right)d\lambda.$$

It follows from these observations that the supremum $\Phi=\sup_{\underset{\sim}{T}(\cdot)}\mathcal{R}(\underset{\sim}{T})$ does not change, if we take it only over the class of pure strategies.

However, *this class may fail to contain an optimal strategy*; in other words, there may not exist a pure strategy that attains the supremum in (3.18)). This can be seen readily by considering *continuous, strictly decreasing* functions $\underline{M}_i(\cdot), i=1,\dots d$, say with $\underline{M}_i(0)=1$. Then the inverses $\sigma_i(\cdot)=\underline{M}_i^{-1}(\cdot)$, and their sum $\tau(m)=\sum_{i=1}^d\sigma_i(m)$, are also continuous and strictly decreasing; the same is true of the inverse $N(\cdot)=\tau^{-1}(\cdot)=\left(\sum_{i=1}^d\underline{M}_i^{-1}(\cdot)\right)^{-1}$. Consequently, the optimal strategy

$$T_i^*(t)=\sigma_i(N(t))=\underline{M}_i^{-1}(N(t));\ 0\le t<\infty,\quad i=1,\dots,d$$

engages all projects at all times, and satisfies (3.32) in the stronger form

$$\underline{M}_i(T_i^*(t))=N(t),\quad \forall\ 0\le t<\infty,\quad i=1,\dots,d.$$

3.15 Remark: The construction of (3.26), (3.29) can be modified in many ways, to produce allocation strategies that satisfy (3.20). More precisely, consider nonnegative, measurable functions $P_i(u,\underset{\sim}{\xi}),\ 0\le u<\infty,\ \underset{\sim}{\xi}=(\xi_1,\dots,\xi_d)\in[0,\infty)^d\backslash\{\underset{\sim}{0}\}$ such that, for every $\underset{\sim}{\xi}$, we have:

$$\left\{\begin{array}{c} u\mapsto P_i(u,\underset{\sim}{\xi})\ \text{ is increasing, }\ 0\le P_i(u,\underset{\sim}{\xi})\le\xi_i \\ \text{and }\ \sum_{i=1}^d P_i(u,\underset{\sim}{\xi})=u,\ \text{ for }\ 0\le u\le|\underset{\sim}{\xi}|\stackrel{\triangle}{=}\sum_{i=1}^d\xi_i\end{array}\right\},$$

$\forall\ i=1,\dots,d$. Define $T_i^*(t)$ by (3.26) for $t\notin\mathcal{D}$, and by

$$T_i^*(t)=\sigma_i(m)+P_i(t-\tau(m),\ -\Delta\underset{\sim}{\sigma}(m));\qquad \tau(m)\le t<\tau(m-),\ m=N(t) \tag{3.37}$$

for $t \in \mathcal{D}$. Here, $\Delta \underset{\sim}{\sigma}(m) = \{\Delta\sigma_i(m)\}_{i=1}^d = \{\sigma_i(m) - \sigma_i(m-)\}_{i=1}^d$.

Now (3.26), (3.37) clearly define an allocation strategy that satisfies (3.25) (hence also (3.20)). For instance, we may take $P_i(u, \underset{\sim}{\xi}) = u\xi_i/|\underset{\sim}{\xi}|,\ 0 \le u \le |\underset{\sim}{\xi}|,$ so that (3.37) becomes

$$T_i^*(t) = \sigma_i(m) + \frac{\sigma_i(m-) - \sigma_i(m)}{\tau(m-) - \tau(m)}(t - \tau(m)), \quad m = N(t). \tag{3.38}$$

One can envision additional such examples.

4. RANDOM FIELDS AND OPTIONAL INCREASING PATHS

Consider a complete probability space $(\Omega, \mathcal{F}, P)$ and a multi-parameter *filtration* (i.e., family of sub-σ-fields of $\mathcal{F}$)

$$\mathbf{F} = \{\mathcal{F}(\underset{\sim}{s})\}_{\underset{\sim}{s} \in \triangle} \tag{4.1}$$

indexed by the elements of the nonnegative cone $\triangle = [0, \infty)^d$ in $\mathcal{R}^d$ with $d \ge 2$. This filtration is assumed to have the following properties:

$$\mathcal{F}(\underset{\sim}{s}) \subseteq \mathcal{F}(\underset{\sim}{r}) \quad \text{for} \quad \underset{\sim}{s} \le \underset{\sim}{r}; \tag{$F.1$}$$

$$\mathcal{F}(\underset{\sim}{0}) \quad \text{contains all the negligible sets in } \mathcal{F}; \tag{$F.2$}$$

$$\mathcal{F}(\underset{\sim}{s}) = \bigcap_{\substack{\underset{\sim}{r} \in \triangle \\ \underset{\sim}{s} < \underset{\sim}{r}}} \mathcal{F}(\underset{\sim}{r}), \quad \forall \ \underset{\sim}{s} \in \triangle; \quad \text{and} \tag{$F.3$}$$

$$\left\{ \begin{matrix} \mathcal{F}(\underset{\sim}{s}), \mathcal{F}(\underset{\sim}{r}) \ \text{ are conditionally independent} \\ \text{given } \ \mathcal{F}(\underset{\sim}{s} \wedge \underset{\sim}{r}), \ \text{ for any } \ (\underset{\sim}{s}, \underset{\sim}{r}) \in \triangle^2. \end{matrix} \right\} \tag{$F.4$}$$

We are denoting here by $\underset{\sim}{s} \le \underset{\sim}{r} \Leftrightarrow_{def} s_i \le r_i, \ \forall\ i = 1, \ldots, d$ the usual partial ordering in $\triangle$, whilst $\underset{\sim}{s} < \underset{\sim}{r} \Leftrightarrow_{def} \underset{\sim}{s} \le \underset{\sim}{r}, \underset{\sim}{s} \ne \underset{\sim}{r}$. Conditions (F.1) – (F.3) are extensions of the one-parameter theory's "usual conditions". Condition (F.4) is standard in multi-parameter theory (cf. Cairoli & Walsh (1975), Walsh (1976/77)), where it seems to be essential for the development of multi-parameter martingales; it implies $\mathcal{F}(\underset{\sim}{s}) \cap \mathcal{F}(\underset{\sim}{r}) = \mathcal{F}(\underset{\sim}{s} \wedge \underset{\sim}{r}),\ \forall (\underset{\sim}{s}, \underset{\sim}{r}) \in \triangle^2$. Condition (F.4) is obviously satisfied in the special case

$$(4.2) \qquad \left\{ \begin{array}{c} \mathcal{F}(\underset{\sim}{s}) = \bigvee_{i=1}^d \mathcal{F}_i(s_i),\ \underset{\sim}{s} = (s_1, \ldots, s_d) \in \triangle, \quad \text{with} \\ \textit{independent} \text{ one-parameter filtrations } \ \mathbf{F}_i = \{\mathcal{F}_i(u)\}_{u \geq 0}, i = 1, \ldots, d. \end{array} \right\}$$

(For an example of a multi-parameter filtration $\mathbf{F}$ which satisfies the condition (F.4) but cannot be written in the form of (4.2), see Remark 4.8 below.)

Let us denote by $\mathcal{S}$ the class of *stopping points* of the filtration $\mathbf{F}$, i.e., of random variables $\underset{\sim}{\nu} : \Omega \to [0, \infty]^d$ that satisfy $\{\underset{\sim}{\nu} \leq \underset{\sim}{s}\} \in \mathcal{F}(\underset{\sim}{s}), \forall\, \underset{\sim}{s} \in \triangle$. And for every $\underset{\sim}{\nu} \in \mathcal{S}$, let

$$\mathcal{F}(\underset{\sim}{\nu}) \stackrel{\triangle}{=} \{A \in \mathcal{F} \,/\, A \cap \{\underset{\sim}{\nu} \leq \underset{\sim}{s}\} \in \mathcal{F}(\underset{\sim}{s}), \quad \forall\ \underset{\sim}{s} \in \triangle\}$$

be the σ-field of events associated with the stopping point $\underset{\sim}{\nu}$. For any pair $\underset{\sim}{\sigma}, \underset{\sim}{\nu}$ of stopping points in $\mathcal{S}$, we have $\{\underset{\sim}{\sigma} \leq \underset{\sim}{\nu}\} \in \mathcal{F}(\underset{\sim}{\nu})$, as well as $\mathcal{F}(\underset{\sim}{\sigma}) \subseteq \mathcal{F}(\underset{\sim}{\nu})$ if $\underset{\sim}{\sigma} \leq \underset{\sim}{\nu}$.

A random field $\mathbf{X} : \triangle \times \Omega \to \mathcal{R}$ with $E|X(\underset{\sim}{s})| < \infty,\ \forall\ \underset{\sim}{s} \in \triangle$ is called an $\mathbf{F}$ - *supermartingale*, if it is $\mathbf{F}$-adapted (that is, $X(\underset{\sim}{s})$ is $\mathcal{F}(\underset{\sim}{s})$-measurable, $\forall \underset{\sim}{s} \in \Delta$), and

$$(4.3) \qquad E[X(\underset{\sim}{r})|\mathcal{F}(\underset{\sim}{s})] \leq X(\underset{\sim}{s}) \ \text{ a.s.}, \quad \forall\ \underset{\sim}{s} \leq \underset{\sim}{r}\,;$$

$\mathbf{X}$ is called an $\mathbf{F}$ - *martingale* if both $\mathbf{X}$ and $-\mathbf{X}$ are $\mathbf{F}$ - supermartingales. For any $\mathbf{F}$ - supermartingale $\mathbf{X}$ with right-continuous paths, and any two stopping points $\underset{\sim}{\sigma} \leq \underset{\sim}{\nu}$ in $\mathcal{S}$, we have the analogue

$$(4.4) \qquad E[X(\underset{\sim}{\nu})|\mathcal{F}(\underset{\sim}{\sigma})] \leq X(\underset{\sim}{\sigma}), \quad \text{a.s.}$$

of the optional sampling theorem, provided that *one* of the following conditions holds:

$$(4.5)(i) \qquad \underset{\sim}{\sigma},\ \underset{\sim}{\nu} \quad \text{are a.s. bounded, or}$$

$$(4.5)(ii) \qquad \mathbf{X} \ \text{ is bounded from below, or}$$

(4.5)(iii) the family $\{X(\underset{\sim}{\nu})\}_{\underset{\sim}{\nu}\in\mathcal{S}}$ is uniformly integrable.

4.1 Proposition: *For a random field* $\mathbf{X} : \triangle \times \Omega \to \mathbf{R}$*, the following are equivalent:*

(a) $\mathbf{X}$ *is an* $\mathbf{F}$ *- supermartingale:*

(b) $\mathbf{X}$ *is* $\mathbf{F}$ *- adapted, integrable, and* $E[X(\underset{\sim}{r})|\mathcal{F}(\underset{\sim}{s})] \le X(\underset{\sim}{r} \wedge \underset{\sim}{s})$ *a.s.,* $\forall\ \underset{\sim}{r}, \underset{\sim}{s}$ *in* $\triangle$*;*

(c) $\{X(\underset{\sim}{s} + t\underset{\sim}{e}_i), \mathcal{G}^i(\underset{\sim}{s} + t\underset{\sim}{e}_i); 0 \le t < \infty\}$ *is a supermartingale,* $\forall \underset{\sim}{s} \in \triangle, \forall i = 1, \ldots, d.$

◇

Here $\underset{\sim}{e}_i$ is the i^{th} unit vector in $\triangle$, and we have set

(4.6) $$\mathcal{G}^i(\underset{\sim}{r}) \overset{\triangle}{=} \mathcal{F}(\infty, \ldots, r_i, \ldots, \infty) = \sigma\Bigg(\bigcup_{\substack{\underset{\sim}{s}\in\triangle \\ s_i \le r_i}} \mathcal{F}(\underset{\sim}{s})\Bigg), \quad \underset{\sim}{r} = (r_1, \ldots, r_d) \in \Delta.$$

Notice that $\mathcal{F}(\underset{\sim}{r}) = \bigcap_{i=1}^{r} \mathcal{G}^i(\underset{\sim}{r}), \quad \forall\ \underset{\sim}{r} \in \Delta.$

4.2 Definition: A family $\underset{\sim}{T} = \{\underset{\sim}{T}(t), 0 \le t < \infty\}$ of $\triangle$-valued random variables is called an *optional increasing path* (O.I.P.) for $\underset{\sim}{s} = (s_1, \ldots, s_d) \in \triangle$, if

(i) $T_i(\cdot)$ is increasing, with $T_i(0) = s_i, \ \forall i = 1, \ldots, d;$

(ii) $\sum_{i=1}^{d}(T_i(t) - s_i) = t, \ \forall\ 0 \le t < \infty$ almost surely; and if

(iii) $\{\underset{\sim}{T}(t) \le \underset{\sim}{r}\} = \bigcap_{i=1}^{d}\{T_i(t) \le r_i\} \in \mathcal{F}(\underset{\sim}{r}), \ \forall\ \underset{\sim}{r} \in \triangle, \ 0 \le t < \infty.$

We shall denote by $\mathcal{A}(\underset{\sim}{s})$ the class of all such O.I.P.'s. ◇

Conditions (i), (ii) are (almost sure) generalizations of the concept of allocation strategy in Definition 3.3; in particular, each component $T_i(\cdot)$ of an optional increasing path is absolutely continuous with respect to Lebesgue measure, and we have the analogue

(4.7) $$T_i(t) = s_i + \int_0^t \chi_i(u)du; \qquad 0 \le t < \infty, \quad i = 1, \ldots, d$$

of (3.16), for suitable measurable processes $\chi_i : [0, \infty) \times \Omega \to [0, 1]$ that satisfy (3.17), almost surely. Condition (iii) implies that each random vector $\underset{\sim}{T}(t)$ $(0 \le t < \infty)$ is an $\mathbf{F}$ -

stopping point, and each random variable $T_i(t)$ $(0 \le t < \infty, i = 1, \ldots, d)$ a stopping time of the one-parameter filtration

$$\mathbf{F}^i = \{\mathcal{F}^i(\theta), \quad 0 \le \theta < \infty\}, \quad \text{where} \quad \mathcal{F}^i(\theta) \overset{\triangle}{=} \mathcal{G}(\theta \underset{\sim}{e}_i) = \mathcal{F}(\infty, \ldots, \theta, \ldots, \infty)$$

as in (4.6). In particular then, for every $\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s})$,

$$\mathbf{F}(\underset{\sim}{T}) = \{\mathcal{F}(\underset{\sim}{T}(t)), \quad 0 \le t < \infty\} \tag{4.9}$$

is a *one-parameter filtration* that satisfies the usual conditions, and the following is a consequence of the optional sampling theorem.

4.3 Proposition: *Let* $\mathbf{X} : \triangle \times \Omega \to \mathcal{R}$ *be an* $\mathbf{F}$ *– (super) martingale with right-continuous paths, and* $\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s})$. *Then*

$$\mathbf{X}(\underset{\sim}{T}) = \{X(\underset{\sim}{T}(t)), \quad 0 \le t < \infty\} \tag{4.10}$$

is a right-continuous, $\mathbf{F}(\underset{\sim}{T})$ *– local (super)martingale.*

4.4 Remark: The adjective "local" can be deleted from the conclusion of Proposition 4.3, if condition (4.5)(iii) holds (or, in the case of a supermartingale $\mathbf{X}$, if (4.5)(ii) holds).

4.5 Proposition: *It* $\underset{\sim}{T}$ *is an optional increasing path in* $\mathcal{A}(\underset{\sim}{0})$, *and* $\sigma : \Omega \to [0, \infty]$ *is a stopping time of the one-parameter filtration* $\mathbf{F}(\underset{\sim}{T})$ *in (4.9), then* $\underset{\sim}{T}(\sigma)$ *is a stopping point of the multi-parameter filtration* $\mathbf{F}$ *of (4.1); in particular, every* $T_i(\sigma), i = 1, \ldots, d$ *is then a stopping time of the one-parameter filtration* $\mathbf{F}^i$ *in (4.8).*

These results can be found in Walsh (1976/77), Walsh (1981); see also Krengel & Sucheston (1987), Edgar & Sucheston (1992), Cairoli & Dalang (1996). We conclude this section with Proposition 4.7 below, a corollary of Propositions 4.1 and 4.3 which will be very useful in section 7. For each $i = 1, \ldots, d$, let us introduce the one-parameter filtration

$$\mathbf{F}_i = \{\mathcal{F}_i(\theta), 0 \le \theta < \infty), \qquad \text{where} \quad \mathcal{F}_i(\theta) \overset{\triangle}{=} \mathcal{F}(\theta \underset{\sim}{e}_i) = \mathcal{F}(0, \ldots, \theta, \ldots, 0). \tag{4.11}$$

The filtrations $\{\mathbf{F}_i\}_{i=1}^d$ satisfy the usual conditions. We shall assume throughout that these filtrations are quasi-left-continuous.

It is useful to keep in mind here the interpretation of the σ-algebra $\mathcal{F}_i(\theta)$ of (4.11) (respectively, $\mathcal{F}^i(\theta)$ of (4.8)) as representing the "history of the i^{th} project during the time-interval $[0, \theta]$, with this project considered in isolation from all the rest" (respectively, "with the i^{th} project viewed in a context where the future evolution of all other projects is known in advance at $t = 0$").

The following is a crucial *enlargement of filtration* property, which links the "small" filtration $\mathbf{F}_i = \{\mathcal{F}_i(\theta), 0 \le \theta < \infty\}$ to the "large" filtration $\mathbf{F}^i = \{\mathcal{F}^i(\theta), 0 \le \theta < \infty\}$, under the condition (F.4).

4.6 Lemma: *Every $\mathbf{F}_i$ - (super) martingale is also an $\mathbf{F}^i$ - (super) martingale.*

Proof: It suffices to verify that, for every $0 \le u < t < \infty$ and for every bounded, $\mathcal{F}_i(t)$-measurable random variable ξ, we have

$$\mathbf{E}[\xi|\mathcal{F}_i(u)] = \mathbf{E}[\xi|\mathcal{F}^i(u)] \quad (= \mathbf{E}[\xi|\mathcal{F}^i(u) \vee \mathcal{F}_i(u)])$$

almost surely. In other words, it suffices to check that

$$\mathcal{F}_i(t), \mathcal{F}^i(u) \ \text{ are conditionally independent, given } \mathcal{F}_i(u).$$

But this follows directly from (F.4) which, with $\underset{\sim}{s} = (0, \ldots, t, \ldots, 0)$ and $\underset{\sim}{r} = (\infty, \ldots, u, \ldots, \infty)$, postulates that $\mathcal{F}_i(t) = \mathcal{F}(\underset{\sim}{s})$ and $\mathcal{F}^i(u) = \mathcal{F}(\underset{\sim}{r})$ are conditionally independent, given $\mathcal{F}(\underset{\sim}{s} \wedge \underset{\sim}{r}) = \mathcal{F}(0, \ldots, u, \ldots, 0) = \mathcal{F}_i(u)$. $\blacksquare$

4.7 Proposition: *For a fixed $\underset{\sim}{s} = (s_1, \ldots, s_d) \in \triangle$, let $\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s})$ be an optional increasing path and, for every $i = 1, \ldots, d$, let $Q_i = \{Q_i(u), s_i \le u < \infty\}$ be an $\mathbf{F}_i$ – martingale* *(respectively, supermartingale of class D) with RCLL paths. Then the process*

$$X(t) = \sum_{i=1}^{d} \int_0^t e^{-\alpha(\theta - T_i(\theta))} dQ_i(T_i(\theta)), \quad 0 \le t < \infty$$

is an $\mathbf{F}(\underset{\sim}{T})$ *– local (super) martingale.*

Proof: For every $i = 1, \ldots, d$, the process

$$K_i(\theta) \stackrel{\Delta}{=} \int_{s_i}^{\theta} e^{\alpha\xi} dQ_i(\xi) = e^{\alpha\theta} Q_i(\theta) - e^{\alpha s_i} Q_i(s_i) - \alpha \int_{s_i}^{\theta} e^{\alpha\xi} Q_i(\xi) d\xi; \quad s_i \le \theta < \infty$$

is an $\mathbf{F}_i$ – martingale (respectively, supermartingale) with RCLL paths. Thus from Lemma 4.6 and Proposition 4.1 (implication (c) $\Rightarrow$ (a)), the random field $K(\underset{\sim}{r}) \stackrel{\Delta}{=} \sum_{i=1}^d K_i(r_i)$, $\underset{\sim}{r} \in \triangle$, $\underset{\sim}{s} \le \underset{\sim}{r}$ is an $\mathbf{F}$ – martingale (respectively, supermartingale). It develops then (from Proposition 4.3) that $K(\underset{\sim}{T}(t)) = \sum_{i=1}^d K_i(T_i(t)) = \sum_{i=1}^d \int_0^t e^{\alpha T_i(\theta)} dQ_i(T_i(\theta))$, $0 \le t < \infty$ is an $\mathbf{F}(\underset{\sim}{T})$ – local (super)martingale, hence the same is true of $X(t) = \int_0^t e^{-\alpha\theta} dK(\underset{\sim}{T}(\theta))$, $0 \le t < \infty$. $\diamond$

4.8 Remark: Let us denote by $\mathcal{B}(\triangle)$ the class of Borel sets of $\triangle = [0,\infty)^d$, and by $\mathcal{W} : \mathcal{B}(\triangle) \times \Omega \to \mathcal{R}$ the *white-noise* (set-valued) *process* corresponding to Lebesque measure λ on $\mathcal{B}(\triangle)$; cf. Walsh (1986), Chapter 1. This is a family of Gaussian random variables $\{W(A)\}_{\substack{A \in \mathcal{B}(\triangle) \\ \lambda(A) < \infty}}$ with

(i) $\mathbf{E}W(A) = 0, \quad \mathbf{E}(W(A))^2 = \lambda(A),$ as well as

(ii) $W(A), W(B)$ independent, and $W(A \cup B) = W(A) + W(B)$ (a.s.), for any disjoint sets A, B in $\mathcal{B}(\triangle)$.

It is easy to check that this zero-mean Gaussian family has covariance function $\mathbf{E}[W(A)W(B)] = \lambda(A \cap B)$.

Now consider the multi-parameter process $B(\underset{\sim}{r}) \stackrel{\triangle}{=} W((\underset{\sim}{0}, \underset{\sim}{r}])$, $\underset{\sim}{r} \in \triangle$, called *Brownian sheet* on $\triangle$, and its associated filtration $\mathbf{F}$ as in (4.1) with

$$\mathcal{F}(\underset{\sim}{r}) \stackrel{\triangle}{=} \sigma(B(\underset{\sim}{s}) / \underset{\sim}{s} \le \underset{\sim}{r}) = \sigma(W(A) / A \in \mathcal{B}(\triangle), A \subseteq (\underset{\sim}{0}, \underset{\sim}{r}])$$

where we have set $(\underset{\sim}{0}, \underset{\sim}{r}] \equiv \times_{i=1}^{d}(0, r_i] \quad \text{for } \underset{\sim}{r} \in \triangle$.

It is not hard to see that this filtration $\mathbf{F}$ satisfies the condition (F.4), but cannot be written in the form (4.2); we owe this observation to Prof. Robert Dalang.

5. GENERAL DYNAMIC ALLOCATION PROBLEM

For every $i \in \{1, ..., d\}$ consider a positive, $\mathbf{F}_i$ – progressively measurable process $H_i = \{h_i(\theta),\ 0 \leq \theta < \infty\}$ such that

$$E \int_0^\infty e^{-\alpha\theta} h_i(\theta) d\theta < \infty, \tag{5.1}$$

where $\alpha \in (0, \infty)$ is a given "discount factor". With these ingredients, and for any given $\underset{\sim}{s} \in \triangle$, consider now the "total discounted reward"

$$\mathcal{R}(\underset{\sim}{T}) = \sum_{i=1}^{d} \int_0^\infty e^{-\alpha t} h_i(T_i(t)) dT_i(t) \tag{5.2}$$

as in (3.8), corresponding to an optional increasing path $\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s})$. The general *Dynamic Allocation* (or "Multi-Armed Bandit") *Problem* is to maximize the conditional expectation $E[\mathcal{R}(\underset{\sim}{T})|\mathcal{F}(\underset{\sim}{s})]$ over $\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s})$, and to find an optional increasing path $\underset{\sim}{T}^* \in \mathcal{A}(\underset{\sim}{s})$ that attains the supremum

$$\Phi(\underset{\sim}{s}) \stackrel{\triangle}{=} \operatorname{esssup}_{\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s})} E[\mathcal{R}(\underset{\sim}{T})|\mathcal{F}(\underset{\sim}{s})]. \tag{5.3}$$

The random variable of (5.3) is called the *value* of, and any optional increasing path $\underset{\sim}{T}^*$ with $\Phi(\underset{\sim}{s}) = E[\mathcal{R}(\underset{\sim}{T}^*)|\mathcal{F}(\underset{\sim}{s})]$ a.s. is called *optimal* for, the Dynamic Allocation Problem at $\underset{\sim}{s} \in \triangle$.

5.1 Interpretation: Suppose there are d projects, competing simultaneously for the attention of a decision-maker; this latter has at his disposal the choice of an *allocation strategy*, modelled here by an optional increasing path $\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{0})$ with the Interpretation

3.5 (we are taking $\underset{\sim}{s} = \underset{\sim}{0}$ and $\mathcal{F}(\underset{\sim}{0}) = \{\emptyset, \Omega\}$ *mod.P*, for concreteness). Engaging the i^{th} project at the calendar time t yields an instantaneous reward $h_i(T_i(t))$, multiplied by the intensity $\chi_i(t) = \frac{dT_i(t)}{dt}$, $0 \le \chi_i(t) \le 1$ of the engagement, and discounted at the rate $\alpha > 0$; thus, the total discounted reward corresponding to an allocation strategy (optional increasing path) $\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{0})$ is given by $\mathcal{R}(\underset{\sim}{T})$ of (5.2), and the value of the problem by $\Phi(\underset{\sim}{0}) = \sup_{\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{0})} E\mathcal{R}(\underset{\sim}{T})$ as in (5.3). If the rewards $h_i(\cdot)$ are deterministic, $\underset{\sim}{T}(\cdot)$ simply has to satisfy the requirements of Definition 3.3. If, however, the rewards are random, the definition of a strategy $\underset{\sim}{T}$ has to reflect the fact that

$$\left\{ \begin{array}{c} \text{decisions about which project(s) to engage, and with what intensity, have to be} \\ \text{made in a non–anticipative way, based on the “histories” } \mathcal{F}_i(r_i) \text{ of the various} \\ \text{projects (accumulated up to their respective engagement times } r_i,\ i = 1, ..., d) \\ \text{by the calendar time } |\underset{\sim}{r}| = \sum_{i=1}^d r_i. \end{array} \right\}$$

This *non-anticipativity* requirement is captured by condition (iii) of Definition 4.2. For certain purposes it is useful to embed the problem of (5.3) into a more general Stochastic Control problem, in which the decision-maker has the additional option of “retiring” (i.e., abandoning all projects) and receiving a lump-sum reward $m \ge 0$.

5.2 Definition: *Mixed Dynamic Allocation/Stopping Problem.* For any given $\underset{\sim}{s} \in \triangle$ and $m \in [0, \infty)$, maximize the conditional expectation $E[\mathcal{R}(\underset{\sim}{T}, \rho; m)|\mathcal{F}(\underset{\sim}{s})]$ of the total discounted reward

$$\mathcal{R}(\underset{\sim}{T}, \rho; m) \triangleq \sum_{i=1}^{d} \int_0^{\rho} e^{-\alpha t} h_i(T_i(t)) dT_i(t) + m e^{-\alpha\rho} \tag{5.4}$$

over the class of “admissible policies”

$$\mathcal{P}(\underset{\sim}{s}) \triangleq \{(\underset{\sim}{T}, \rho) \,/\, \underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s}),\ \rho \text{ is an } \mathbf{F} - \text{stopping time}\}. \qquad \diamond \tag{5.5}$$

The value of this problem will be denoted by

$$\Phi(\underset{\sim}{s}; m) \triangleq \operatorname{esssup}_{(\underset{\sim}{T}, \rho) \in \mathcal{P}(\underset{\sim}{s})} E[\mathcal{R}(\underset{\sim}{T}, \rho; m)|\mathcal{F}(\underset{\sim}{s})], \tag{5.6}$$

and clearly it satisfies $\Phi(\underset{\sim}{s}) \equiv \Phi(\underset{\sim}{s};0)$, as well as

$$
\begin{aligned}
\text{(5.7)} \qquad \Phi(\underset{\sim}{s};m) - m &= \operatorname{esssup}_{(\underset{\sim}{T},\rho)\in\mathcal{P}(\underset{\sim}{s})} E\Big[\sum_{i=1}^{d} \int_0^{\rho} e^{-\alpha t}\{h_i(T_i(t)) - \alpha m\} dT_i(t) \,\Big|\, \mathcal{F}(\underset{\sim}{s})\Big] \\
&= \operatorname{esssup}_{(\underset{\sim}{T},\rho)\in\mathcal{P}(\underset{\sim}{s})} E[\mathcal{R}(\underset{\sim}{T},\rho;m) - m|\mathcal{F}(\underset{\sim}{s})].
\end{aligned}
$$

6. THE ONE–PARAMETER PROBLEM

The analogue of (5.6), (5.7) in the one-parameter case ($d = 1$) is the *optimal stopping problem* with value

(6.1)
$$
\begin{aligned}
V_i(u;m) &\triangleq \operatorname{esssup}_{\sigma\in\mathcal{S}_i(u)} E\Big[\int_u^{\sigma} e^{-\alpha(\theta-u)} h_i(\theta) d\theta + m e^{-\alpha(\sigma-u)} \,\Big|\, \mathcal{F}_i(u)\Big] \\
&= m + \operatorname{esssup}_{\sigma\in\mathcal{S}_i(u)} E\Big[\int_u^{\sigma} e^{-\alpha(\theta-u)}\{h_i(\theta) - \alpha m\} d\theta \,\Big|\, \mathcal{F}_i(u)\Big]; \quad 0 \le u, m < \infty
\end{aligned}
$$

and $V_i(\infty;m) \equiv m$, for fixed $i \in \{1,...,d\}$. Here $\mathcal{S}_i(u)$ is the set of stopping times σ of the filtration $\mathbf{F}_i$ in (4.11), with values in $[u,\infty]$. From the standard theory of Optimal Stopping, it is well–known (e.g. Neveu (1975), Ch. VI, El Karoui (1981), or Karatzas (1993)) that for every given $m \in [0,\infty)$, the process

$$
\begin{aligned}
\text{(6.2)} \qquad Q_i(u;m) &\triangleq e^{-\alpha u} V_i(u;m) + \int_0^u e^{-\alpha\theta} h_i(\theta) d\theta \\
&= m + e^{-\alpha u}[V_i(u;m) - m] + \int_0^u e^{-\alpha\theta}\{h_i(\theta) - \alpha m\} d\theta, \quad 0 \le u < \infty
\end{aligned}
$$

is a regular $\mathbf{F}_i$ – supermartingale of class D, with RCLL paths; it is the smallest $\mathbf{F}_i$-supermartingale which dominates the continuous, positive process $me^{-\alpha u} + \int_0^u e^{-\alpha\theta} h_i(\theta) d\theta$, $0 \le u < \infty$. Furthermore, the stopping time

$$\text{(6.3)} \qquad \sigma_i(u;m) \triangleq \inf\{\theta \ge u / V_i(\theta;m) = m\} \text{ attains the supremum in (6.1);}$$

and the stopped process

$$\text{(6.4)} \qquad Q_i(\theta \wedge \sigma_i(u;m);m), \quad u \le \theta < \infty \quad \text{is an } \mathbf{F}_i \text{ – martingale.}$$

On the other hand, for every fixed $u \in [0, \infty)$, the process $m \mapsto \sigma_i(u; m)$ is decreasing and right–continuous with values in $[u, \infty]$, $\sigma_i(u; 0) = \infty$ and $\sigma_i(u; \infty) = u$, almost surely; and the mapping $m \mapsto V_i(u; m)$ is convex and increasing, with values in $[m, \infty)$, $\lim_{m \to \infty}[V_i(u; m) - m] = 0$ and right–hand derivative

$$\frac{\partial^+}{\partial m} V_i(u; m) \triangleq \lim_{\delta \downarrow 0} \frac{1}{\delta}[V_i(u; m + \delta) - V_i(u; m)] = E[e^{-\alpha(\sigma_i(u;m) - u)} | \mathcal{F}_i(u)] \tag{6.5}$$

almost surely. Equivalently, in integral form,

$$V_i(u; m) - m = E\Big[\int_m^\infty (1 - e^{-\alpha(\sigma_i(u;\lambda) - u)}) d\lambda \;\Big|\; \mathcal{F}_i(u)\Big], \quad a.s. \tag{6.6}$$

All these results are proved in section 2 of El Karoui & Karatzas (1994).

Consider now, for fixed $u \in [0, \infty)$, the decreasing, positive, and $\mathbf{F}_i$ – adapted process

$$\underline{M}_i(u, \theta) \triangleq \inf\{m \geq 0 / \sigma_i(u; m) \leq \theta\}, \quad u \leq \theta < \infty, \tag{6.7}$$

the right-continuous inverse of the mapping $m \mapsto \sigma_i(u; m)$, and observe from (3.4) that

$$M_i(u) \triangleq \underline{M}_i(u, u) = \lim_{\theta \downarrow u} \underline{M}_i(u, \theta) \tag{6.8}$$

satisfies $\{m > 0 / V_i(u; m) = m\} = [M_i(u), \infty)$, modulo P. In other words, $M_i(u)$ has the significance of *equitable surrender value* at $\sigma = u$ *for the* i^{th} *project viewed in isolation*, as it is the smallest value of the parameter $m \geq 0$ that makes immediate stopping profitable at $\sigma = u$ in (6.1). We call the $\mathbf{F}_i$ – progressively measurable process $M_i(u), 0 \leq u < \infty$ the *Gittins index process* for the i^{th} project. Its lower envelope coincides with the process of (6.7): $\underline{M}_i(u, \theta) = \inf_{u \leq s \leq \theta} M_i(s), \quad u \leq \theta < \infty$. The index also admits the so-called *forwards induction* interpretation

$$M_i(u) = \operatorname{esssup}_{\substack{\sigma > u \\ \sigma \in \mathcal{S}_i(u)}} \frac{E[\int_u^\sigma e^{-\alpha\theta} h_i(\theta) d\theta | \mathcal{F}_i(u)]}{E[\int_u^\sigma e^{-\alpha\theta} d\theta | \mathcal{F}_i(u)]}, \quad a.s. \tag{6.9}$$

This index process $M_i(\cdot)$ of (6.8), (6.9) will not play a role in the theory of section 7.

6.1 Proposition: *For every* $i \in \{1, ..., d\}$, $0 \le s$, $m < \infty$ *we have the a.s. identities*

$$V_i(\sigma_i(s;m);m) = m, \quad V_i(\sigma_i(s;m-);m) = m, \tag{6.10}$$

$$e^{-\alpha s} V_i(s;m) = E\Big[\int_s^\infty \alpha e^{-\alpha\theta}(m \vee \underline{M}_i(s,\theta))d\theta \;\Big|\; \mathcal{F}_i(s)\Big]; \tag{6.11}$$

in particular,

$$e^{-\alpha s} V_i(s;0) = E\Big[\int_s^\infty e^{-\alpha\theta} h_i(\theta)d\theta \;\Big|\; \mathcal{F}_i(s)\Big] = E\Big[\int_s^\infty e^{-\alpha\theta} \alpha \underline{M}_i(s,\theta)d\theta \;\Big|\; \mathcal{F}_i(s)\Big]. \tag{6.12}$$

6.2 Proposition: *For every* $i \in \{1, ..., d\}$ *and* $s_i \in [0, \infty)$, *the process*

(6.13)

$$\mathcal{U}_i(\theta) := e^{-\alpha\theta}[V_i(\theta; \underline{M}_i(s_i,\theta)) - \underline{M}_i(s_i,\theta)] + \int_{s_i}^\theta e^{-\alpha u}\{h_i(u) - \alpha \underline{M}_i(s_i,u)\}du, \quad s_i \le \theta < \infty$$

is an $\mathbf{F}_i$ – *martingale with RCLL paths.*

For proofs of the properties (6.9), (6.11)–(6.13), including a detailed discussion and examples, see section 3 of El Karoui & Karatzas (1994).

Proof of (6.10): To simplify typography, let us drop the subscript i throughout, and write $\sigma_s(m)$ for $\sigma(s;m)$. Clearly $V(\sigma_s(m);m) = m$ a.s., from the definition (6.3) and the right-continuity of $V(\cdot;m)$. With $0 < \lambda < m$ we have $\sigma_s(m) \le \sigma_s(m-) \le \sigma_s(\lambda)$, whence

$$e^{-\alpha s}[V(s;\lambda) - \lambda] = E\Big[\int_s^{\sigma_s(\lambda)} e^{-\alpha\theta}(h(\theta) - \alpha\lambda)d\theta \;\Big|\; \mathcal{F}(s)\Big], \; a.s. \tag{6.14}$$

and from (6.4):

$$E[Q(\sigma_s(m-);\lambda)|\mathcal{F}(\sigma_s(m))] = Q(\sigma_s(m);\lambda), \; a.s. \tag{6.15}$$

Letting $\lambda \uparrow m$ in (6.14), (6.15) we obtain, respectively,

$$e^{-\alpha s}[V(s;m) - m] = E\Big[\int_s^{\sigma_s(m-)} e^{-\alpha\theta}(h(\theta) - \alpha m)d\theta \;\Big|\; \mathcal{F}(s)\Big] \tag{6.16}$$

and

(6.17)

$$E[e^{-\alpha\sigma_s(m-)}(V(\sigma_s(m-);m)-m)|\mathcal{F}(\sigma_s(m))] = -E\Big[\int_{\sigma_s(m)}^{\sigma_s(m-)} e^{-\alpha\theta}(h(\theta)-\alpha m)d\theta \,\Big|\, \mathcal{F}(\sigma_s(m))\Big],$$

almost surely. Now take expectations in (6.17) and note (by virtue of (6.16), and of (6.14) with λ replaced by m) that the resulting right-hand side is then equal to zero; this means that the expectation of the nonnegative random variable $e^{-\alpha\sigma_s(m-)}[V(\sigma_s(m-);m)-m]$ is also equal to zero, whence $V(\sigma_s(m-);m)=m$, a.s. $\diamond$

6.3 Remark: The representations of (6.5), (6.6), (6.11), (6.12), as well as the properties of (6.2), (6.4), (6.13), *all remain in force if one replaces the filtration* $\mathbf{F}_i$ *by the filtration* $\mathbf{F}^i$ *of (4.8)*; recall Lemma 4.6.

7. OPTIMALITY IN THE GENERAL PROBLEM

For fixed $\underset{\sim}{s}=(s_1,...,s_d)\in\triangle$, consider the Dynamic Allocation Problem of section 5 with the reward processes H_i replaced by the *decreasing, right–continuous processes*

$$h_i'(\theta) \triangleq \alpha \underline{M}_i(s_i,\theta), \quad s_i\le\theta<\infty, \quad i=1,...,d \tag{7.1}$$

(in the notation of (6.3), (6.7)), and denote by

$$\Psi(\underset{\sim}{s}) \triangleq \operatorname{esssup}_{\underset{\sim}{T}\in\mathcal{A}(\underset{\sim}{s})} E[\mathcal{R}'(\underset{\sim}{T})|\mathcal{F}(\underset{\sim}{s})], \ \mathcal{R}'(\underset{\sim}{T}) = \sum_{i=1}^d \int_0^\infty e^{-\alpha t} h_i'(T_i(t))dT_i(t) \tag{7.2}$$

the value of this new problem.

The complete solution of Problem (7.2) is provided, as an easy corollary, by the theory of section 3. One introduces the decreasing, right–continuous processes (inverses of each other)

$$\tau(m;\underset{\sim}{s}) \triangleq \sum_{i=1}^d (\sigma_i(s_i;m)-s_i), 0\le m<\infty \quad \text{and} \quad N(t;\underset{\sim}{s}) \triangleq \inf\{m\ge 0/\tau(m;\underset{\sim}{s})\le t\} \tag{7.3}$$

for $0 \le t < \infty$ and constructs, *for every fixed* $\omega \in \Omega$, the allocation strategy $\underset{\sim}{T}^*(\cdot,\omega) - \underset{\sim}{s} = (T_1^*(\cdot,\omega) - s_1, ..., T_d^*(\cdot,\omega) - s_d)$ as in Proposition 3.9 with $\tau(\cdot)$, $N(\cdot)$ replaced by $\tau(\cdot,\omega;\underset{\sim}{s})$, $N(\cdot,\omega;\underset{\sim}{s})$ respectively. This defines an optional increasing path $\underset{\sim}{T}^* \in \mathcal{A}(\underset{\sim}{s})$ which satisfies *pathwise*:

(i) The "synchronization identity" (3.20), in the form

$$\sum_{i=1}^{d}([T_i^*(t) - s_i] \wedge [\sigma_i(s_i;\lambda-) - s_i]) = t \wedge \tau(\lambda-;\underset{\sim}{s}), \quad \forall\, 0 \le t < \infty, \quad 0 < \lambda < \infty. \tag{7.4}$$

(ii) The "index–type property" (3.35) in the form

$$\sum_{i=1}^{d}\int_0^\infty 1_{\{\underline{M}_i(s_i,T_i^*(t)) < \max_{1\le j\le d}\underline{M}_j(s_j,T_j^*(t))\}}dT_i^*(t) = 0; \tag{7.5}$$

this states that "every $T_i^*(\cdot)$ grows only (modulo Lebesgue measure) on the set $\{t \ge 0 / \underline{M}_i(s_i,T_i^*(t)) = \max_{i\le j\le d}\underline{M}_j(s_j,T_j^*(t))\}$ where the lower–envelope $\underline{M}_i(s_i,T_i^*(\cdot)) = \inf_{s_i\le u\le T_i^*(\cdot)} M_i(u)$ of its index process is maximal among all projects".

(iii) The "dual optimality property" (3.31)–(3.32)

$$\underline{\mathcal{M}}(\underset{\sim}{s},\underset{\sim}{T}(t)) \ge N(t,\underset{\sim}{s}) = \underline{\mathcal{M}}(\underset{\sim}{s};\underset{\sim}{T}^*(t)), \; 0 \le t < \infty; \qquad \forall\, \underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s}) \tag{7.6}$$

with the notation $\underline{\mathcal{M}}(\underset{\sim}{s},\underset{\sim}{r}) \triangleq \max_{1\le i\le d}\underline{M}_i(s_i,r_i)$ for $\underset{\sim}{r} \in \triangle$, $\underset{\sim}{s} \le \underset{\sim}{r}$.

(iv) The *optimality property (3.22)*, in the form

$$\mathcal{R}'(\underset{\sim}{T}) \le \mathcal{R}'(\underset{\sim}{T}^*) = \int_0^\infty \alpha e^{-\alpha t}N(t;\underset{\sim}{s})dt = \int_0^\infty [1 - e^{-\alpha\tau(\lambda;\underset{\sim}{s})}]d\lambda, \; \forall\, \underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s}). \tag{7.7}$$

In particular, $\underset{\sim}{T}^*$ is optimal for the problem of (7.2):

$$\begin{aligned} E[\mathcal{R}'(\underset{\sim}{T})|\mathcal{F}(\underset{\sim}{s})] \le E[\mathcal{R}'(\underset{\sim}{T}^*)|\mathcal{F}(\underset{\sim}{s})] = \Psi(\underset{\sim}{s}) &= E\Big[\int_0^\infty \alpha e^{-\alpha t}N(t;\underset{\sim}{s})dt \;\Big|\; \mathcal{F}(\underset{\sim}{s})\Big] \\ &= E\Big[\int_0^\infty (1 - e^{-\alpha\tau(\lambda;\underset{\sim}{s})})d\lambda \;\Big|\; \mathcal{F}(\underset{\sim}{s})\Big] \quad a.s, \quad \forall\, \underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s}). \end{aligned} \tag{7.8}$$

It should be stressed that *(7.7), (7.8) hold without any assumption on the multi-parameter filtration* $\mathbf{F}$; in particular, the assumption (F.4) is not needed for these properties. Now the fundamental result of this work, Theorem 7.1 below, states that $\underset{\sim}{T}^*$ is optimal also in our original, general Dynamic Allocation Problem (5.4). For this theorem, the condition (F.4), on the multi-parameter filtration $\mathbf{F}$, is crucial; the result was established in our earlier work El Karoui & Karatzas (1994) under the independent assumption (4.2) on $\mathbf{F}_i$.

7.1 Theorem: *For every fixed* $\underset{\sim}{s} \in \triangle$, *the general Dynamic Allocation Problem (5.3) with general reward proceses* $\{h_i(t), t \geq s_i\}_{i=1}^d$, *has the same value, namely*

(7.9)
$$\Phi(\underset{\sim}{s}) \equiv \Psi(\underset{\sim}{s}) = E\Big[\int_0^\infty \alpha e^{-\alpha t} N(t;\underset{\sim}{s})dt \,\Big|\, \mathcal{F}(\underset{\sim}{s})\Big] = E\Big[\int_0^\infty (1 - e^{-\alpha\tau(\lambda;\underset{\sim}{s})})d\lambda \,\Big|\, \mathcal{F}(\underset{\sim}{s})\Big], a.s.$$

and the same optimal strategy $\underset{\sim}{T}^* \in \mathcal{A}(\underset{\sim}{s})$, *namely*

$$E[\mathcal{R}(\underset{\sim}{T})|\mathcal{F}(\underset{\sim}{s})] \leq E[\mathcal{R}(\underset{\sim}{T}^*)|\mathcal{F}(\underset{\sim}{s})] = \Phi(\underset{\sim}{s}) \;\; a.s., \quad \forall\, \underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s}), \tag{7.10}$$

as the problem (7.2) with the decreasing rewards $\{h_i'(t) = \alpha \underline{M}_i(s_i,t), t \geq s_i\}_{i=1}^d$ *of (7.1).* $\diamond$

The remainder of the section will be devoted to the proof of this result. Clearly, it suffices to show that

$$\left\{ \begin{array}{c} K(t) \triangleq \sum_{i=1}^d \int_0^t e^{-\alpha(\theta - T_i^*(\theta))} d\mathcal{U}_i(T_i^*(\theta)), \quad 0 \leq t < \infty \\ \text{is a uniformly integrable } \; \mathbf{F}(\underset{\sim}{T}^*) - \text{martingale} \end{array} \right\} \tag{7.11}$$

in the notation of (6.13), that

$$\left\{ \begin{array}{c} K(\infty) = \Lambda(\infty), \quad \text{almost surely, where} \\ \Lambda(t) \triangleq \sum_{i=1}^d \int_0^t e^{-\alpha\theta}[h_i(T_i^*(\theta)) - \alpha \underline{M}_i(s_i, T_i^*(\theta))]dT_i^*(\theta), \quad 0 \leq t < \infty \end{array} \right\} \tag{7.12}$$

and that the positive process

$$\left\{ \begin{array}{c} Z(t;\underset{\sim}{T}) \triangleq e^{-\alpha t}\Psi(\underset{\sim}{T}(t)) + \sum_{i=1}^d \int_0^t e^{-\alpha\theta} h_i(T_i(\theta))dT_i(\theta), \quad 0 \leq t < \infty \\ \text{is an } \; \mathbf{F}(\underset{\sim}{T}) - \text{supermartingale}, \quad \forall\, \underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s}). \end{array} \right\} \tag{7.13}$$

Because then we have from (7.13)

$$\Psi(\underset{\sim}{s}) = Z(0;\underset{\sim}{T}) \geq E[Z(\infty;\underset{\sim}{T})|\mathcal{F}(\underset{\sim}{s})] = E[\mathcal{R}(\underset{\sim}{T})|\mathcal{F}(\underset{\sim}{s})] \;\; a.s., \quad \forall\, \underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s}), \tag{7.14}$$

whereas $E[\Lambda(\infty)|\mathcal{F}(\underset{\sim}{s})] = E[K(\infty)|\mathcal{F}(\underset{\sim}{s})] = K(0) = 0$ from (7.11), (7.12) gives, with the notation of (5.2), (5.3), (7.2):

$$E[\mathcal{R}(\underset{\sim}{T}^*)|\mathcal{F}(\underset{\sim}{s})] = E[\mathcal{R}'(\underset{\sim}{T}^*)|\mathcal{F}(\underset{\sim}{s})] = \Psi(\underset{\sim}{s}). \tag{7.15}$$

All the claims of Theorem 7.1 follow from (7.14), (7.15).

To establish (7.11) - (7.13) we shall need a few auxiliary Lemmata. For results on the stochastic calculus of one–parameter processes see, for instance, Karatzas & Shreve (1991) or Revuz & Yor (1991).

7.2 Lemma: *For any $\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s})$, $m \geq 0$, the process*

$$X(t) \triangleq \sum_{i=1}^{d} \int_0^t e^{-\alpha(\theta - T_i(\theta))} d_\theta Q_i(T_i(\theta); m), \quad 0 \leq t < \infty \tag{7.16}$$

in the notation of (6.2), is an $\mathbf{F}(\underset{\sim}{T})$–local supermartingale.

Proof: The processes $Q_i(\cdot; m)$ of (6.2) are $\mathbf{F}_i$ – supermartingales of class D with RCLL paths, so the result follows from Proposition 4.7.

7.3 Lemma: *For any $\underset{\sim}{T} \in \mathcal{A}(\underset{\sim}{s})$, the process*

$$Y(t) \triangleq e^{-\alpha t}\Psi(\underset{\sim}{T}(t)) - \sum_{i=1}^{d} \int_0^t e^{-\alpha(\theta - T_i(\theta))} d_\theta (e^{-\alpha T_i(\theta)} V_i(T_i(\theta); 0)), \quad 0 \leq t < \infty \tag{7.17}$$

is an $\mathbf{F}(\underset{\sim}{T})$–supermartingale.

Proof: From (6.6), (7.9) and Remark 6.3:

$$e^{-\alpha T_i(t)}[V_i(T_i(t); m) - m] = E\left[\int_m^\infty \left\{ e^{-\alpha T_i(t)} - e^{-\alpha \sigma_i(T_i(t);\lambda)} \right\} d\lambda \,\middle|\, \mathcal{F}^i(T_i(t))\right],$$

$$e^{-\alpha t}\Psi(\underset{\sim}{T}(t)) = E\left[\int_0^\infty \left\{ e^{-\alpha \sum_{j=1}^d (T_j(t) - s_j)} - e^{-\alpha \sum_{j=1}^d (\sigma_j(T_j(t);\lambda) - s_j)} \right\} d\lambda \,\middle|\, \mathcal{F}(\underset{\sim}{T}(t))\right], a.s.$$

Therefore, taking $\underset{\sim}{s} = \underset{\sim}{0}$ to simplify typography, we have to show for $u < t$:

$$E\left[\sum_{i=1}^{d}\int_0^\infty \int_u^t e^{-\alpha\sum_{j\neq i}T_j(\theta)} d_\theta\left(e^{-\alpha T_i(\theta)} - e^{-\alpha\sigma_i(T_i(\theta);\lambda)}\right)d\lambda \,\middle|\, \mathcal{F}(\underset{\sim}{T}(u))\right]$$
$$\geq E\left[\int_0^\infty\left\{\left(e^{-\alpha\sum_{j=1}^d T_j(t)} - e^{-\alpha\sum_{j=1}^d \sigma_j(T_j(t);\lambda)}\right)\right.\right.$$
$$\left.\left. - \left(e^{-\alpha\sum_{j=1}^d T_j(u)} - e^{-\alpha\sum_{j=1}^d \sigma_j(T_j(u);\lambda)}\right)\right\}d\lambda \,\middle|\, \mathcal{F}(\underset{\sim}{T}(u))\right],$$

or equivalently:

$$E\left[\int_0^\infty\left\{e^{-\alpha\sum_{j=1}^d \sigma_j(T_j(t);\lambda)} - e^{-\alpha\sum_{j=1}^d \sigma_j(T_j(u);\lambda)}\right.\right.$$
$$\left.\left. - \sum_{i=1}^d\int_u^t e^{-\alpha\sum_{j\neq i}T_j(\theta)} d_\theta(e^{-\alpha\sigma_i(T_i(\theta);\lambda)})\right\}d\lambda \,\middle|\, \mathcal{F}(\underset{\sim}{T}(u))\right]$$
$$\geq E\left[\int_0^\infty\left\{e^{-\alpha\sum_{j=1}^d T_j(t)} - e^{-\alpha\sum_{j=1}^d T_j(u)}\right.\right.$$
$$\left.\left. - \sum_{i=1}^d\int_u^t e^{-\alpha\sum_{j\neq i}T_j(\theta)} d_\theta(e^{-\alpha T_i(\theta)})\right\}d\lambda \,\middle|\, \mathcal{F}(\underset{\sim}{T}(u))\right], \tag{7.18}$$

almost surely. From Lemma 7.4 below and the continuity of the $T_j(\cdot)$'s, the right–hand–side of (7.18) is zero. On the other hand, in the notation of Lemma 7.4 and with $A_j(\cdot) \equiv \alpha\sigma_j(T_j(\cdot);\lambda), j = 1,\ldots,d$, the expression inside the braces on the left-hand-side of (7.18) dominates

$$e^{-A(t)} - e^{-A(u)} - \sum_{i=1}^d \int_u^t e^{-A^{(i)}(\theta-)} d(e^{-A_i(\theta)})$$

which is nonnegative (again thanks to Lemma 7.4), and the result is proved.

7.4 Lemma: *Let $A_i : [0,\infty) \to [0,\infty)$ be increasing, right–continuous functions, and set $A = \sum_{j=1}^d A_j$, $A^{(i)} = A - A_i = \sum_{j\neq i} A_j$ for $i = 1,...,d$. Then*

$$e^{-A(t)} - e^{-A(0)} - \sum_{i=1}^d \int_0^t e^{-A^{(i)}(s-)} d(e^{-A_i(s)}) =$$
$$= \sum_{s\leq t} e^{-A(s-)}\left[(e^{-\sum_{i=1}^d \Delta A_i(s)} - 1) - \sum_{i=1}^d (e^{-\Delta A_i(s)} - 1)\right], \quad 0 \leq t < \infty \tag{7.19}$$

is an increasing, pure jump function (identically equal to zero, if A is continuous).

Proof: Let us recall the change-of-variable formula for an increasing, right–continuous function C (e.g. Revuz & Yor (1991), p.6), namely

$$F(C(t)) = F(C(0)) + \int_0^t F'(C(s-))dC(s) + \sum_{s\le t}\{F(C(s)) - F(C(s-)) - F'(C(s-))\Delta C(s)\}$$

where $\Delta C(s) \triangleq C(s) - C(s-)$. We shall apply this formula to the function $F(x) = e^{-x}$, twice: first with $C = A$,

$$e^{-A(t)} - e^{-A(0)} + \sum_{i=1}^d \int_0^t e^{-A(s-)} dA_i(s) = \sum_{s\le t} e^{-A(s-)}[e^{-\Delta A(s)} - 1 + \Delta A(s)] \tag{7.20}$$

and then with $C = A_i$,

$$e^{-A_i(t)} - e^{-A_i(0)} + \int_0^t e^{-A_i(s-)} dA_i(s) = \sum_{s\le t} e^{-A_i(s-)}[e^{-\Delta A_i(s)} - 1 + \Delta A_i(s)].$$

It follows from this last expression that

$$\int_0^t e^{-A^{(i)}(s-)} d(e^{-A_i(s)}) + \int_0^t e^{-A(s-)} dA_i(s) = \sum_{s\le t} e^{-A(s-)}[e^{-\Delta A_i(s)} - 1 + \Delta A_i(s)]; \tag{7.21}$$

summing up over $i = 1, ..., d$ in (7.21), and substituting the resulting expression for $\sum_{i=1}^d \int_0^t e^{-A(s-)} dA_i(s)$ into (7.20), we obtain (7.19). Now the right–hand side of (7.19) is an increasing (pure jump) function, since: $1 - e^{-\sum_{i=1}^d \Delta A_i(s)} \le \sum_{i=1}^d (1 - e^{-\Delta A_i(s)})$.

Proof of (7.13): From Lemmata 7.3 and 7.2 (with $m = 0$) we see that

$$Z(t; \underset{\sim}{T}) = X(t) + Y(t) = e^{-\alpha t}\Psi(\underset{\sim}{T}(t)) + \sum_{i=1}^d \int_0^t e^{-\alpha\theta} h_i(T_i(\theta)) dT_i(\theta), \quad 0 \le t < \infty$$

is an $\mathbf{F}(\underset{\sim}{T})$–local supermartingale; because it is positive, it is actually an $\mathbf{F}(\underset{\sim}{T})$ – supermartingale.

7.5 Lemma: *For every $i = 1, \ldots, d$ we have, in the notation of sections 3 and 6:*

$$Y_i(t) \triangleq V_i(T_i^*(t); \underline{M}_i(s_i, T_i^*(t)) - \underline{M}_i(s_i, T_i^*(t)) = 0$$

on the set $([0,\infty)\backslash\mathcal{D})\bigcup(\bigcup_{m\in\mathbf{D}}([\tau(m),\tau(m-))\backslash(y_{i-1}(m),y_i(m))))$.

Proof: To simplify typography we take $\underset{\sim}{s}=\underset{\sim}{0}$ and write $\sigma_i(m)$ for $\sigma_i(0;m),\tau(m)$ for $\tau_i(m;\underset{\sim}{0})$, $\underline{M}_i(u)$ for $\underline{M}_i(0,u)$, and $N(t)$ for $N(t;\underset{\sim}{0})$.

For $t\notin\mathcal{D}$, we have $T_i^*(t)=\sigma_i(m-)$ from (3.26), with $m=N(t)$. For $\tau(m)\le t<\tau(m-)$ with some $m\in\mathbf{D}$, we have from (3.29): $T_i^*(t)=\left\{\begin{array}{ll}\sigma_i(m-); & \tau(m)\le t\le y_{i-1}(m)\\ \sigma_i(m); & y_i(m)\le t<\tau(m-)\end{array}\right\}$ if $m\in\mathbf{D}_i$ (i.e., if $\sigma_i(\cdot)$ has a jump at m), and $T_i^*(t)=\sigma_i(m)=\sigma_i(m-)$ if $m\notin\mathbf{D}_i$ (i.e., if $\sigma_i(\cdot)$ is continuous at m); here again, $m=N(t)$. We distinguish the following cases:

$\underline{(a)\ m\notin\mathcal{B}_i}$. Then $\underline{M}_i(\sigma_i(m-))=\underline{M}_i(\sigma_i(m))=m$, from (3.5), (3.7), and $T_i^*(t)=\sigma_i(m\pm)$; thus $V_i(T_i^*(t);\underline{M}_i(T_i^*(t)))=V_i(\sigma_i(m\pm);\underline{M}_i(\sigma_i(m\pm)))=V_i(\sigma_i(m\pm);m)=m=\underline{M}_i(T_i^*(t))$, from (6.10).

$\underline{(b)\ m\in\mathcal{B}_i, m\in\mathbf{D}_i}$. Then $\underline{M}_i(\sigma_i(m-))<\underline{M}_i(\sigma_i(m))=m$, and there exists a point $r=\sigma_i(m-)\in\mathbf{B}_i$ such that $m=\underline{M}_i(r-)$.

If $T_i^*(t)=\sigma_i(m)$, then $\underline{M}_i(T_i^*(t))=m$ and thus $V_i(T_i^*(t);\underline{M}_i(T_i^*(t)))=V_i(\sigma_i(m);m)$ $=m=\underline{M}_i(T_i^*(t))$, from (6.10).

If $T_i^*(t)=\sigma_i(m-)$, then $T_i^*(t)=r=\sigma_i(\underline{M}_i(r))$ so that, again from (6.10),

$$V_i(T_i^*(t);\underline{M}_i(T_i^*(t)))=V_i(\sigma_i(\underline{M}_i(r));\underline{M}_i(r))=\underline{M}_i(r)=\underline{M}_i(T_i^*(t)). \tag{7.22}$$

$\underline{(c)\ m\in\mathcal{B}_i, m\notin\mathbf{D}_i}$: Then $\underline{M}_i(\sigma_i(m-))=\underline{M}_i(\sigma_i(m))<m$, and there exists a point $r\stackrel{\triangle}{=}\sigma_i(m-)=\sigma_i(m)\in\mathbf{B}_i$ such that $m\in(\underline{M}_i(r),\underline{M}_i(r-)]$. In this case $T_i^*(t)=r=\sigma_i(\underline{M}_i(r))$, and (6.10) leads again to (7.22).

Proof of (7.12) *(due to Prof. D.L. Ocone):* To simplify notation, let us take $\underset{\sim}{s}=\underset{\sim}{0}$ again. From (6.13), the definition of $K(\cdot),\Lambda(\cdot)$ and the notation of Lemma 7.5, the equality

$K(\infty) = \Lambda(\infty)$ is equivalent to

$$\sum_{i=1}^{d} \int_0^\infty e^{-\alpha(t - T_i^*(t))} d_t \Big(e^{-\alpha T_i^*(t)} Y_i(t) \Big) = 0. \tag{7.23}$$

From Lemma 7.5 and (3.29), the left-hand-side of (7.23) equals

$$\begin{aligned} &\sum_{m \in \mathbf{D}} \sum_{i=1}^{d} \int_{y_{i-1}(m)}^{y_i(m)} e^{-\alpha(t - T_i^*(t))} d_t (e^{-\alpha T_i^*(t)} Y_i(t)) \\ &= \sum_{i=1}^{d} \sum_{m \in \mathcal{D}_i} e^{-\alpha(y_{i-1}(m) - \sigma_i(m))} \cdot \Big(e^{-\alpha T_i^*(t)} Y_i(t) \Big) \Big|_{t=y_{i-1}(m)}^{t=y_i(m)} \\ &= \sum_{i=1}^{d} \sum_{m \in \mathcal{D}_i} e^{-\alpha y_{i-1}(m)} [Y_i(y_i(m)) e^{\alpha \Delta \sigma_i(m)} - Y_i(y_{i-1}(m))]. \end{aligned} \tag{7.24}$$

Now (7.23) follows from Lemma 7.5, since $Y_i(y_i(m)) = Y_i(y_{i-1}(m)) = 0, \;\; \forall \; i = 1, \ldots, d$ (for $i = d$, observe that $y_d(m) = \tau(m-) \notin \mathcal{D}$).

7.6 Remark: Notice that the explicit form (3.29) of the strategy $\underset{\sim}{T}^*(\cdot)$, in particular the fact that each $T_i^*(\cdot)$ grows linearly and at slope $+1$ on intervals of the type $(y_{i-1}(m), y_i(m)), m \in \mathbf{D}_i$, has been used in a crucial way in the proof of (7.12). This property is *not* shared by the allocation strategies of Remark 3.15, for instance by (3.38).

Proof of (7.11): The processes $\mathcal{U}_i(\cdot)$ of (6.13) are $\mathbf{F}_i$-martingales with RCLL paths; thus, Proposition 4.7 shows that $K(t) = \sum_{i=1}^d \int_0^t e^{-\alpha(\theta - T_i^*(\theta))} d\mathcal{U}_i(T_i^*(\theta)), \quad 0 \le t < \infty$ is an $\mathbf{F}(\underset{\sim}{T}^*)$-local martingale. To prove the claim of (7.11), it suffices to show that

$$\text{the family } \;\; \{K(\tau)\}_{\tau \in \mathcal{S}(\mathbf{F}(\underset{\sim}{T}^*))} \;\; \text{is uniformly integrable,} \tag{7.25}$$

where $\mathcal{S}(\mathbf{F}(\underset{\sim}{T}^*))$ is the class of stopping times $\tau : \Omega \to [0, \infty]$ of the filtration $\mathbf{F}(\underset{\sim}{T}^*)$.

Now in the notation of Lemma 7.5 and (7.12), we have the decomposition $K(\cdot) = \Lambda(\cdot) + W(\cdot)$, where

$$W(t) \triangleq \sum_{i=1}^{d} \int_0^t e^{-\alpha(\theta - T_i^*(\theta))} d_\theta (e^{-\alpha T_i^*(\theta)} Y_i(\theta)), \quad 0 \le t < \infty. \tag{7.26}$$

The process $\Lambda(\cdot)$ of (7.12) is dominated a.s. by the random variable

$$L \triangleq \sum_{i=1}^{d} e^{\alpha s_i} \Big\{ \int_{s_i}^{\infty} e^{-\alpha u} h_i(u) du + \int_{s_i}^{\infty} \alpha e^{-\alpha u} \underline{M}_i(s_i, u) du \Big\}$$

which is integrable, thanks to $E \int_0^\infty \alpha e^{-\alpha\theta} \underline{M}_i(s_i, \theta) d\theta = E \int_0^\infty e^{-\alpha\theta} h_i(\theta) d\theta < \infty$ from (6.12) and (5.1). Therefore, in order to prove (7.25), it suffices to show that

$$\text{the family } \{W(\tau)\}_{\tau \in \mathcal{S}(\mathbf{F}(\underset{\sim}{T}^*))} \text{ is uniformly integrable.} \tag{7.27}$$

To carry this out let us suppose again, for simplicity of notation, that $\underset{\sim}{s} = \underset{\sim}{0}$, and observe that $W(\tau)$ of (7.26) can be written equivalently as

$$\begin{aligned} W(\tau) = & \sum_{\substack{m \in \mathbf{D} \\ m > N(\tau)}} \sum_{i=1}^{d} \int_{y_{i-1}(m)}^{y_i(m)} e^{-\alpha(\theta - T_i^*(\theta))} d_\theta \Big(e^{-\alpha T_i^*(\theta)} Y_i(\theta) \Big) \\ & + \sum_{i=1}^{d} \int_{y_{i-1}(m)}^{\tau} e^{-\alpha(\theta - T_i^*(\theta))} d_\theta \Big(e^{-\alpha T_i^*(\theta)} Y_i(\theta) \Big) \Big|_{m = N(\tau)} \end{aligned} \tag{7.28}$$

by analogy with (7.24). From (7.24) and the discussion that follows it, the first (double) summation on the right-hand-side of (7.28) is zero, and the second equals

$$\begin{aligned} W(\tau) &= \int_{y_{k-1}(m)}^{\tau} e^{-\alpha(y_{k-1}(m) - \sigma_k(m))} d_\theta \Big(e^{-\alpha T_k^*(\theta)} Y_k(\theta) \Big) \\ &= e^{-\alpha(T_k^*(\tau) + y_{k-1}(m) - \sigma_k(m))} Y_k(\tau) = e^{-\alpha\tau} \Big[V_k(\sigma; \lambda) - \lambda \Big] \Big|_{\substack{\sigma = T_k^*(\tau) \\ \lambda = \underline{M}_k(\sigma)}} \\ &\le e^{-\alpha\sigma} \cdot \sup_{0 < \lambda < \infty} \Big[V_k(\sigma; \lambda) - \lambda \Big] \Big|_{\sigma = T_k^*(\tau)} \quad , a.s. \end{aligned}$$

in the notation of (3.29) with $m = N(\tau),\ k = k(m),\ y_{k-1}(m) \le \tau < y_k(m)$. In particular, the process $W(\cdot)$ is nonnegative, and in order to prove (7.27) it suffices to show (by virtue of Proposition 4.5) that

$$\begin{gathered} \text{the family } \{\sup\nolimits_{0<m<\infty} (e^{-\alpha\sigma} [V_i(\sigma; m) - m])\}_{\sigma \in \mathcal{S}(\mathbf{F}^i)} \\ \text{is uniformly integrable}, \quad \forall \;\; i = 1, \ldots, d, \end{gathered} \tag{7.29}$$

where $\mathcal{S}(\mathbf{F}^i)$ is the class of stopping times $\sigma : \Omega \to [0,\infty]$ of the filtration $\mathbf{F}^i$ in (4.8). But from (6.11), (6.12) and Remark 6.3 we obtain, for every $0 \le s < \infty$ and $m \in (0,\infty)$:

$$\begin{aligned} e^{-\alpha s}[V_i(s;m) - m] &= E\left[\int_s^\infty \alpha e^{-\alpha\theta}\Big(\underline{M}_i(s,\theta) - m\Big)^+ d\theta \,\middle|\, \mathcal{F}^i(s)\right] \\ &\le E\left[\int_s^\infty \alpha e^{-\alpha\theta}\underline{M}_i(s,\theta)d\theta \,\middle|\, \mathcal{F}^i(s)\right] = E\left[\int_s^\infty \alpha e^{-\alpha\theta} h_i(\theta)d\theta \,\middle|\, \mathcal{F}^i(s)\right] \\ &\le E[A_i|\mathcal{F}^i(s)], \quad A_i \triangleq \int_0^\infty \alpha e^{-\alpha\theta} h_i(\theta)d\theta, \end{aligned}$$

whence

$$\sup_{0<m<\infty}\left(e^{-\alpha\sigma}[V_i(\sigma;m) - m]\right) \le E[A_i|\mathcal{F}^i(\sigma)], \quad a.s.$$

for every $\sigma \in \mathcal{S}(\mathbf{F}^i)$. The assumption $E(A_i) < \infty$ of (5.1) implies the uniform integrability of the family $\{E[A_i|\mathcal{F}^i(\sigma)]\}_{\sigma\in\mathcal{S}(\mathbf{F}^i)}$, and (7.29) follows.

7.7 Remark: The methods of this section also solve the mixed Dynamic Allocation/ Stopping Problem of Definition 5.2: for any given $m \ge 0$ and $\underset{\sim}{s} \in \triangle$, the value of (5.6) is given by

$$\begin{aligned} \Phi(\underset{\sim}{s};m) &= E\left[\int_0^\infty \alpha e^{-\alpha t}(N(t;\underset{\sim}{s}) \vee m)dt \,\middle|\, \mathcal{F}(\underset{\sim}{s})\right] \\ &= m + E\left[\int_m^\infty (1 - e^{-\alpha\tau(\lambda;\underset{\sim}{s})})d\lambda \,\middle|\, \mathcal{F}(\underset{\sim}{s})\right], \quad a.s. \end{aligned} \tag{7.30}$$

and the supremum of (5.6) is attained by the pair $(\underset{\sim}{T}^*, \rho^*) \in \mathcal{P}(\underset{\sim}{s})$, with $\rho^* = \tau(m;\underset{\sim}{s})$ and $\underset{\sim}{T}^* \in \mathcal{A}(\underset{\sim}{s})$ the same as in Theorem 7.1. To check that $\tau(m;\underset{\sim}{s})$ *is indeed an* $\mathbf{F}(\underset{\sim}{T}^*)$ – *stopping time*, for any $m \ge 0$ and $\underset{\sim}{s} \in \triangle$, observe that we have from (7.4):

$$\{\tau(m;\underset{\sim}{s}) \le t\} = \bigcap_{i=1}^d \{\sigma_i(s_i;m) \le T_i^*(t)\} = \{\underset{\sim}{\sigma}(\underset{\sim}{s};m) \le \underset{\sim}{T}^*(t)\} \in \mathcal{F}(\underset{\sim}{T}^*(t)), \quad \forall\, t \ge 0 \tag{7.31}$$

for every $0 \le t < \infty$. This is because both $\underset{\sim}{\sigma}(\underset{\sim}{s};m) \equiv (\sigma_1(s_1;m),...,\sigma_d(s_d;m))$ and $\underset{\sim}{T}^*(t) = (T_1^*(t),...,T_d^*(t))$ are stopping points of the multi-parameter filtration $\mathbf{F}$ – the former from Definition 4.2(iii), and the latter thanks to

$$\{\underset{\sim}{\sigma}(\underset{\sim}{s};m) \le \underset{\sim}{r}\} = \bigcap_{i=1}^{d} \{\sigma_i; m) \le r_i\} = \bigcap_{i=1}^{d} \{\underline{M}_i(s_i;r_i) \le m\} \in \bigcap_{i=1}^{d} \mathcal{F}^i(r_i) = \mathcal{F}(\underset{\sim}{r})$$

which is valid for every $\underset{\sim}{r} \in \triangle,\ \underset{\sim}{s} \le \underset{\sim}{r}$.

8. THE CASE OF INDEPENDENT "BROWNIAN" BANDITS

As a simple illustration of the results in sections 3 and 7, let us consider the situation of the Problem in (5.3) with $\underset{\sim}{s} = \underset{\sim}{0}$ and reward processes

$$h_i(\theta) = \eta_i(W_i(\theta)), \quad 0 \le \theta < \infty, \quad i = 1, ..., d. \tag{8.1}$$

Here each $\eta_i : \mathcal{R} \to (0,\infty)$ is a strictly increasing function of class $C^1(\mathcal{R})$, and $W_1(\cdot), ..., W_d(\cdot)$ are *independent* standard Brownian motions; we denote by $\mathbf{F}_i$ the augmentation of the natural filtration$\{\mathcal{F}^{W_i}(\theta)\}_{0\le\theta<\infty}$, and define the multi–parameter filtration $\mathbf{F}$ as in (4.2).

In this case, it is well–known (Karatzas (1984); see also Mandelbaum (1987), El Karoui & Karatzas (1994), §3.10) that the index-process $M_i(\cdot)$ of (6.8), (6.9) is $M_i(\theta) = \nu_i(W_i(\theta)),\ 0 \le \theta < \infty$, where the *index function* $\nu_i : \mathcal{R} \to (0,\infty)$ is given as

$$\nu_i(x) = \frac{1}{\alpha} \int_0^\infty \eta_i\Big(x + \frac{z}{\sqrt{2\alpha}}\Big) e^{-z} dz \tag{8.2}$$

and inherits the properties of strict increase and $C^1(\mathcal{R})$ from η_i, for every $i = 1, ..., d$. We shall assume for convenience that all the $\nu_i(\cdot)$'s have the same range, so that their inverses $\nu_i^{-1}(\cdot)$ are all defined on the same domain, and denote by

$$\nu(\cdot) \ \text{ the inverse of } \ \ \nu^{-1}(\cdot) \triangleq \sum_{i=1}^{d} \nu_i^{-1}(\cdot). \tag{8.3}$$

It follows that the decreasing processes $\underline{M}_i(\cdot) \equiv \underline{M}_i(0,\cdot)$ of (6.7) are given as

$$\underline{M}_i(\theta) = \inf_{0\le u\le\theta} M_i(u) = \nu_i(A_i(\theta)), \quad A_i(\theta) \triangleq \min_{0\le u\le\theta} W_i(u), \quad 0 \le \theta < \infty; \tag{8.4}$$

in particular, they are *continuous*, with right–continuous decreasing, pure-jump inverses

$$
\begin{aligned}
\sigma_i(m) &= \inf\{\theta \geq 0 / \underline{M}_i(\theta) \leq m\} \\
&= \inf\{\theta \geq 0 / \nu_i(W_i(\theta)) \leq m\}, \quad 0 \leq m < \infty, \quad i = 1, ..., d.
\end{aligned} \tag{8.5}
$$

Note that, if $\mathbf{D}_i$ stands for the set of jump–points of $\sigma_i(\cdot)$, then

$$
\left\{ \begin{array}{c} \mathcal{D}_i^0 \triangleq \bigcup_{m \in \mathbf{D}_i} (\sigma_i(m), \sigma_i(m-)) \text{ is the set of excursion intervals} \\ \text{away from the origin, of } \nu_i(W_i(\theta)) - \nu_i(A_i(\theta)), \text{ or equivalently of} \\ \textit{the reflecting Brownian motion process } \; R_i(\theta) \triangleq W_i(\theta) - A_i(\theta), 0 \leq \theta < \infty. \end{array} \right\} \tag{8.6}
$$

Now the pure-jump processes $m \mapsto \sigma_i(m)$ $i = 1, ..., d$ are independent, thus the sets $\mathbf{D}_i$ $(i = 1, ..., d)$ *disjoint*, and their union $\mathbf{D} = \cup_{i=1}^d \mathbf{D}_i$ is the set of jump–points of the decreasing, right continuous, *pure-jump process*

$$
\tau(m) = \sum_{i=1}^d \sigma_i(m), \quad 0 \leq m < \infty
$$

with *continuous* (decreasing) inverse $N(t) = \inf\{m \geq 0 / \tau(m) \leq t\}, \quad 0 \leq t < \infty.$

With these ingredients, it is straightforward to see that the dynamic allocation strategy $\underset{\sim}{T}^*(\cdot) = (T_i^*(\cdot), \ldots, T_d^*(t))$ of Proposition 3.9 takes the form

$$
T_i^*(t) = \sigma_i(N(t)) + (t - \tau(N(t))) 1_{\{N(t) \in \mathbf{D}_i\}}, \quad 0 \leq t < \infty, \quad i = 1, ..., d. \tag{8.7}
$$

Now each $\underline{M}_i(\cdot)$ is continuous, so that $\sigma_i(\cdot)$ has no flat stretches (i.e., $\mathcal{B}_i = \emptyset$), and we obtain from (3.7), (3.36) that

$$
\underline{M}_i(T_i^*(t)) = N(t), \quad 0 \leq t < \infty, \quad i = 1, ..., d. \tag{8.8}
$$

In other words, $\underset{\sim}{T}^*(\cdot)$ maintains equal lower envelopes of indices for all projects, at all times, just as in Remark 3.14. From Theorem 7.1, this strategy $\underset{\sim}{T}^*(\cdot)$ attains the supremum of expected discounted reward

$$
\Phi(\underset{\sim}{0}) \triangleq \sup_{\underset{\sim}{T}(\cdot) \in \mathcal{A}(\underset{\sim}{0})} E \sum_{i=1}^d \int_0^\infty e^{-\alpha t} \eta_i(W_i(T_i(t))) dT_i(t). \tag{8.9}
$$

Consider now the nonnegative processes

$$\begin{aligned} \mathcal{S}_i(t) &\triangleq R_i(T_i^*(t)) = (W_i - A_i)(T_i^*(t)) \\ &= \nu_i^{-1}(M_i(T_i^*(t))) - \nu_i^{-1}(N(t)), \quad 0 \le t < \infty, \quad i = 1, ..., d \end{aligned} \tag{8.10}$$

and notice that

$$\mathcal{S}_i(t) > 0 \iff T_i^*(t) \in \mathcal{D}_i^0 \;\Rightarrow\; N(t) \in \mathbf{D}_i.$$

In particular, *at most one of the* $\mathcal{S}_i(t),\ i = 1, ..., d$ *can be strictly positive at any given time* $t > 0$, and we have

$$\int_0^\infty 1_{\{\mathcal{S}_i(t)>0\}} dN(t) = 0 \;\; (i = 1, \dots, d), \quad \text{whence} \quad \int_0^\infty 1_{\{\mathcal{S}(t)>0\}} dN(t) = 0 \tag{8.11}$$

for the sum $\mathcal{S}(t) \triangleq \sum_{i=1}^d \mathcal{S}_i(t)$ of the processes in (8.10). This can be written as

$$\mathcal{S}(t) = W(t) + L(t) \tag{8.12}$$

where $W(t) \triangleq \sum_{i=1}^d W_i(T_i^*(t))$ is standard Brownian motion, and $L(t) \triangleq -\nu^{-1}(N(t))$, in the notation of (8.3), is a continuous, increasing process with $\int_0^\infty 1_{\{\mathcal{S}(t)>0\}} dL(t) = 0$ according to (8.11). From the theory of the Skorohod problem for reflecting Brownian motion (e.g. Karatzas & Shreve (1991), pp.211–212) it develops then that $\mathcal{S}(\cdot)$ is a *reflecting* (standard, one–dimensional) *Brownian motion*, and that

$$L(t) = -\nu^{-1}(N(t)) = \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} \text{ meas}\{0 \le u \le t \,/\, \mathcal{S}(u) \le \epsilon\}$$

is its local time at the origin. The distribution $P[L(t) \in d\ell] = 2(2\pi t)^{-1/2} \exp(-\ell^2/2t) d\ell$, $\ell > 0$ of this random variable is well known (ibid, eqs. (3.6.28) and (2.8.3)), and leads in conjuction with (7.9) and

$$N(t) = \nu(-L(t)) = \Big(\sum_{i=1}^d \nu_i^{-1}\Big)^{-1}(-L(t)), \quad 0 \le t < \infty \tag{8.14}$$

to the computation

$$\Phi(\underset{\sim}{0}) = E \int_0^\infty \alpha e^{-\alpha t} N(t) dt = \int_0^\infty \alpha e^{-\alpha t} E[\nu(-L(t))] dt = \int_0^\infty \nu\Big(-\frac{z}{\sqrt{2\alpha}}\Big) e^{-z} dz \tag{8.15}$$

for the value of the problem (8.9).

8.1 The case d=2, and the "skew Brownian motion": The special case $d = 2$ is studied in detail in section 9 of El Karoui & Karatzas (1994). Here, let us consider the difference

$$X(t) = \mathcal{S}_1(t) - \mathcal{S}_2(t) = B(t) + V(t), \quad 0 \le t < \infty \tag{8.16}$$

of the processes in (8.10), where $B(t) \triangleq W_1(T_1^*(t)) - W_2(T_2^*(t))$ is Brownian motion and

$$V(t) \triangleq (\nu_2^{-1} - \nu_1^{-1})(N(t)) = \varphi(L(t)) \tag{8.17}$$

a process of bounded variation; we have set

$$\varphi(\ell) \triangleq (\nu_2^{-1} - \nu_1^{-1})(\nu(-\ell)). \tag{8.18}$$

Because $\mathcal{S}_1(\cdot)\mathcal{S}_2(\cdot) \equiv 0$, we have $\mathcal{S}_1(\cdot) \equiv X^+(\cdot)$, $\mathcal{S}_2(\cdot) \equiv X_1^-(\cdot)$; thus $X(\cdot) = B(\cdot) + V(\cdot)$ is a semimartingale, and $|X(\cdot)| = \mathcal{S}(\cdot)$ is a reflecting Brownian motion with symmetric local time at the origin

$$L^X(t) \triangleq \lim_{\epsilon \downarrow 0} \frac{1}{2\epsilon} \text{ meas}\{0 \le u \le t / |X(u)| \le \epsilon\} = L(t) \tag{8.19}$$

from (8.13). We conclude that $X(\cdot)$ solves the stochastic equation

$$X(t) = \varphi(L^X(t)) + B(t), \quad B(\cdot) = \text{Brownian motion.} \tag{8.20}$$

If $\nu_1 \equiv \nu_2$ $(\eta_1 \equiv \eta_2)$, then $\varphi = 0$ and $X(\cdot)$ is a Brownian motion. If we have $\nu_2(\alpha x) = \nu_1(x), \forall x \in \mathcal{R}$, for some $0 < \alpha \le 1$, then $\phi(\ell) = \beta\ell$ with $\beta = \frac{1-\alpha}{1+\alpha}$ and $X(\cdot)$ is a so–called *skew Brownian motion* (cf. Harrison & Shepp (1981), Walsh (1978)). In general, the function $\varphi(\cdot)$ of (8.18) is of class $C^1(0,\infty)$, and satisfies $|\varphi'| \le 1$. It might be of independent interest, to develop a general theory for equations of the type (8.20).

ACKNOWLEDGEMENTS: We are deeply grateful to Professor Daniel Ocone for his very thorough reading of an earlier version of this paper, for finding an error in our proof of (7.12), and for providing us with a new argument. The connection (8.16) of Dynamic Allocation with skew-Brownian motion goes back twelve years ago, to A. Mandelbaum (private communication).

REFERENCES

BACCELLI, F., COHEN, G., OLSDER, G.J. & QUADRAT, J.P. (1992) *Synchronization and Linearity: An Algebra for Discrete-Event Systems.* J. Wiley & Sons, New York.

BANKS, J.S. & SUNDARAM, R.K. (1992) Denumerable-armed bandits. *Econometrica* **60**, 1071-1096.

BERGEMANN, D. & VALIMAKI, J. (1993) Learning and strategic pricing. Preprint.

BERRY, D.A. & FRISTEDT, B. (1985) *Bandit Problems: Sequential Allocation of Experiments.* Chapman & Hall, New York.

CAIROLI, R. & DALANG, R.C. (1996) *Sequential Stochastic Optimization.* J. Wiley & Sons, N.Y.

CAIROLI, R. & WALSH, J.B. (1975) Stochastic integrals in the plane. *Acta Mathematica* **134**, 111-183.

DALANG, R.C. (1990) Randomization in the two-armed bandit problem. *Annals of Probability* **18**, 218-225.

EDGAR, G.A. & SUCHESTON, L. (1992) *Stopping Times and Directed Processes.* Encyclopaedia of Mathematics and its Applications, Vol. 47. Cambridge University Press.

EL KAROUI, N. (1981) Les aspects probabilistes du contrôle stochastique. *Lecture Notes in Mathematics* **876**, 73-238. Springer-Verlag, Berlin.

EL KAROUI, N. & KARATZAS, I. (1993) General Gittins index processes in discrete time. *Proceedings of the National Academy of Sciences (USA)* **90**, 1232-1236.

EL KAROUI, N. & KARATZAS, I. (1994) Dynamic allocation problems in continuous time. *Annals of Applied Probability* **4**, 255-286.

EPLETT, W.J.R. (1986) Continuous-time allocation indices and their discrete-time approximations. *Adv. Appl. Probability* **18**, 724-746.

GITTINS, J.C. (1979) Bandit processes and dynamic allocation indices (with discussion). *J. Roy. Statist. Society* **B 41**, 148-164.

GITTINS, J.C. (1989) *Multi-Armed Bandits and Allocation Indices.* J. Wiley & Sons, Chichester.

GITTINS, J.C. & JONES, D.M. (1974) A dynamic allocation index for the sequential design of experiments. *Progress in Statistics* (J. Gani, ed.), 241-266. North-Holland, Amsterdam.

HARRISON, J.M. & SHEPP, L.A. (1981) On skew Brownian motion. *Annals of Probability* **9**, 309-313.

KARATZAS, I. (1984) Gittins indices in the dynamic allocation problem for diffusion processes. *Annals of Probability* **12**, 173-192.

KARATZAS, I. (1993) *Lectures on Optimal Stopping and Stochastic Control.* Department of Statistics, Columbia University, N.Y.

KARATZAS, I. & SHREVE, S.E. (1991) *Brownian Motion and Stochastic Calculus*, Second Edition. Springer-Verlag, New York.

KASPI, H. & MANDELBAUM, A. (1995) Lévy bandits: multi-armed bandits driven by Lévy processes. *Annals of Applied Probability* **5,** 541-565.

KASPI, H. & MANDELBAUM, A. (1996) Multi-armed bandits in discrete time. Preprint.

KRENGEL, V. & SUCHESTON, L. (1981) Stopping rules and tactics for processes indexed by directed sets. *J. Multivariate Anal.* **11**, 199-229.

MANDELBAUM, A. (1986) Discrete multi-armed bandits and multi-parameter processes. *Probab. Theory & Rel. Fields* **71**, 129-147.

MANDELBAUM, A. (1987) Continuous multi-armed bandits and multi-parameter processes. *Annals of Probability* **15**, 1527-1556.

MAZZIOTTO, G. & MILLET, A. (1987) Stochastic control of two-parameter processes, with application to the two-armed bandit problem. *Stochastics* **22**, 251-288.

MENALDI, J.L. & ROBIN, M. (1990) On the optimal reward function of the continuous-time multi-armed bandit problem. *SIAM J. Control & Optimization* **28**, 97-112.

MORIMOTO, H. (1987) Optimal switching for alternating processes. *Appl. Math. Optimization* **16**, 1-17.

NEVEU, J. (1975) *Discrete-Parameter Martingales.* English Translation, North-Holland, Amsterdam and American Elsevier, N.Y.

PRESMAN, E.L. (1990) Poisson version of the two-armed bandit problem with discounting. *Theory Probab. Appl.* **35**, 307-317.

PRESMAN, E.L. & SONIN, I.M. (1983) Two- and many-armed bandit problems with infinite horizon. *Lecture Notes in Mathematics* **1021**, 526-540. Springer-Verlag, Berlin.

PRESMAN, E.L. & SONIN, I.M. (1990) *Sequential Control with Incomplete Data: Bayesian Approach.* English Translation, Academic Press, N.Y.

REVUZ, D. & YOR, M. (1991) *Continuous Martingales and Brownian Motion.* Springer-Verlag, Berlin.

ROTHSCHILD, M. (1974) A two-armed bandit theory of market pricing. *J. Econom. Theory* **9**, 185-202.

TANAKA, T. (1994) The system of quasi-variational inequalities attached to the two-armed bandit problem. *Stochastics* **47**, 139-161.

TSITSIKLIS, J. (1986) A lemma on the multi-armed bandit problem. *IEEE Trans. Autom. Control* **AC-31**, 576-577.

TSITSIKLIS, J. (1994) A short proof of the Gittins index theorem. *Annals of Applied Probability* **4**, 194-199.

VANDERBEI, R.J. (1992) Optimal switching among several Brownian motions. *SIAM J. Control & Optimization* **30**, 1150-1162.

VARAIYA, P., WALRAND, J. & BUYUKKOC, C. (1985) Extensions of the multi-armed bandit problem: the discounted case. *IEEE Trans. Autom. Control* **AC-30**, 426-439.

WALSH, J.B. (1976/77) *Martingales with multi-dimensional parameter and stochastic integrals in the plane.* Lecture Notes, Cours de 3ème Cycle, Université P&M Curie (Paris VI).

WALSH, J.B. (1978) A diffusion with discontinuous local time. *Astérisque* **52-53**, 37-45.

WALSH, J.B. (1981) Optional Increasing Paths. *Lecture Notes in Mathematics* **863**, 172-201. Springer-Verlag, Berlin.

WALSH, J.B. (1986) An Introduction to Stochastic Partial Differential Equations. *Lecture Notes in Mathematics* **1180**, 265-439. Springer-Verlag, Berlin.

WEBER, R.R. (1992) On the Gittins index for multi-armed bandits. *Annals of Applied Probability* **2**, 1024-1033.

WHITTLE, P. (1980) Multi-armed bandits and the Gittins index. *J. Roy. Statist. Society* **B 42**, 143-149.

WHITTLE, P. (1982) *Optimization over Time: Dynamic Programming and Stochastic Control, Volume I.* J. Wiley & Sons, New York.

YUSHKEVICH, A.A. (1988) On the two-armed bandit problem with continuous time parameter and discounted rewards. *Stochastics* **23**, 299-310.