THE OPTIMAL COST EXPANSION OF FINITE CONTROLS FINITE STATES MARKOV CHAINS WITH WEAK AND STRONG INTERACTIONS

F. DELEBECQUE
J.P. QUADRAT

I. Introduction

The purpose of this paper is to study Markov chains with strong and weak transition probabilities called interactions. If we study such Markov chains on a small period of time the weak interactions can be neglegted in first approximation but if we study this process on a time large enough we cannot do this approximation. If we call $0 < \varepsilon << 1$ the order of the weak interactions, we look at the Markov chain on a period of order $\frac{1}{\varepsilon}$ and give for this problem the complete expansion of the expected value of a cost associated to a trajectory. We give also the complete expansion of the optimal cost for the controlled problem. In both cases, we have also the stochastic interpretation of all the term of the expansion. The problem is solved without other hypotheses that : - finite number of states and of values of the control.

We define a fast Markov chain neglegting the transition probabilities of order ε . The introduction of an aggregated Markov chain is necessary to define the expansion. This aggregated Markov chain has for states the final classes of the fast chain and for cost an average cost. This average cost is constant on the final classes of the fast chain its value is the average (relative to the invariant measure of the fast chain) of the initial cost. Its transition probabilities are : for $\bar{x} \neq \bar{x}'$

$$\sum_{x \in \bar{x}} p_{\bar{x}}(x) \, [\sum_{x' \in \bar{x}'} a_{xx'} + \sum_{y \in \bar{y}} a_{xy} q_{\bar{x}'}(y)]$$ where $q_{\bar{x}}(y)$ is the probability

to end in the final class $\bar{x}$ starting from the transient state y, $a_{xx'}$ are the weak interactions $(\bar{x}, \bar{x}')$ are final classes, $\bar{y}$ the set of transient states , of the fast chain, $p_{\bar{x}}$ the invariant measure of the fast chain of support $\bar{x}$. So the definition of this aggregated chain needs only the solving of linear systems, the sizes of which are the numbers of states in the final classes, then the solving of an aggregated problem : computation of the expected value of the aggregated cost on an aggregated trajectory for the aggregated chain gives the first term of the expansion. The other terms of the expansion can be computed by the same decentralized-aggregated way.

This kind of aggregated chains appears in the litterature in Courtois [5] ,Pervorzvanskii-Smirnov [16], Gaistgori-Pervorzvanskii[9], on hypotheses excluding general fast transient chains. These authors study the invariant measure of Markov chains with weak and strong interactions. Here we don't make any hypothesis of this kind and we give the complete expansion, for a ε-discounted cost.

The resolution of the controlled problem needs the introduction of a vector Hamilton-Jacobi-Bellman equation like has done . Veinott [18] in the case of control of Markov chains with small discount rate. This vector H.J.B. equation determines uniquely all the terms of the expansion in the general situation (for the case of a finite number of values for the control). A policy iteration algorithm (Howard's algorithm) gives a way to compute the n-first terms of the expansion. This algorithm needs the solving of linear systems. This can be done by the decentralized aggregated way described above . In the general situation a stochastic interpretation of all the terms of the expansion is given in term of the fast and aggregated chains. In particular the first term of the expansion, in the case where the control does not change the final classes of the fast chain, can be interpreted as the optimal cost of the aggregated chain defined before. In this latter case, the optimization is done simultaneously in the aggregation (definition of the aggregated chain and cost) and in the control of the aggregated chain

The first part gives analytical results and stochastic interpretation for the uncontrolled case, the second one do the same thing for the controlled one.

This work is related to two kinds of litterature :the litterature on the control of finite states Markov chain (for example Bellman [2] , Howard [10] , Derman [8] , Lanery [13] , Veinott[18] , Chîtashvilli [3], Rothblum [17]), and the litterature on perturbation of operator or of Markov chains Kato [11]- Courtois [5], Pervorzvanskii-Smirnov [16], Gaitsgori-Pervorzvanskii[9] . The results obtained are similar to the ones obtained in Chow-Kokotovic [4], for the control of deterministic systems. Applications to management of hydropower systems and a result for controlled diffusion processes are given in Delebecque-Quadrat [6]. The generalized averaging of Bensoussan-Lions-Papanicolaou [1], gives more difficult results for uncontrolled diffusion processes. Schweitzer-Federgruen [19] study a two level-Bellman equation which is similar to the one obtained here with other motivations.

II. MARKOV CHAIN WITH STRONG AND WEAK INTERACTIONS

We study in this part a Markov chain with finite states $x \in E$, we suppose that the number of states card (E) = n, its transition matrix is called M, and its generator M-I is supposed to be equal to $B+\varepsilon A$ where B and A are generators of Markov chains and ε a real number small relatively to $1, 0 < \varepsilon << 1$. This Markov chains is denoted by X_t where t is the time belonging to $\mathbb{N}$.

Given the function

$$f : I \times E \to \mathbb{R}^+ \quad \text{with} \quad \sup_{\substack{x\in E\\ i\in\mathbb{N}}} f_i(x) \leq C_f, \text{ where } C_f \in \mathbb{R}^+$$
$$i \quad x \quad f_i(x)$$

we define :

$$f_\varepsilon : E \to \mathbb{R}^+,$$
$$x \to \sum_{i=0}^{+\infty} \varepsilon^i f_i(x)$$

$\mu>0$ given, we are interested by the expansion in ε of the mapping :

$$(1.1) \qquad V_\varepsilon : \mathbb{E} \to \mathbb{R}, \quad x \to V^\varepsilon_{(x)} = \mathbb{E}^x \sum_{t=0}^{+\infty} \frac{\varepsilon}{(1+\mu\varepsilon)^{t+1}} \; f_\varepsilon \circ X_t$$

But if we denote by ν a random variable which takes its values in $\mathbb{N}$, which is independent of X_t, and with law defined by

$$P(\nu = t) = \frac{\varepsilon\mu}{(1+\mu\varepsilon)^{t+1}} \qquad \text{then :}$$

$$(1.2) \qquad V_\varepsilon = \frac{1}{\mu} \; \mathbb{E} \; f_\varepsilon \circ X_\nu$$

But $\mathbb{E}(\nu) = \frac{1}{\mu\varepsilon}$ and V_ε thanks to (1.2) can be seen as the cost of the

Markov chain on a time scale of order $\frac{1}{\varepsilon}$, and is the solution of Kolmogorov equation :

$$(1.3) \qquad -\mu V_\varepsilon + \frac{1}{\varepsilon} BV_\varepsilon + AV_\varepsilon + f_\varepsilon = 0$$

So the asymptotic study of V_ε defined by (1.1) or (1.2) is also equivalent to the study of the solution of (1.3). In this paragraph we are interested, in the first part, in the expansion of V_ε solution of (1.3), that is the analytical study, and in the second part, in the stochastic interpretation of the terms of this expansion(denoted by $V: \mathbb{N} \times E \to R$ with $i \quad x \quad V_i(x)$

$$V_\varepsilon(x) = \sum_{i=0}^{+\infty} \varepsilon^i V_i(x)).$$

The Markov chain defined on E, of transition matrix B + I which is a stochastic matrix, is called <u>the fast chain</u>, and denoted (Z_t).

We use the notations :

$\mathcal{N}(B)$ is the kernel of the operator B which is $\neq \{0\}$ because B is a generator $B1 = 0$;

$\mathcal{R}(B)$ is the range of the operator B ;

P is the spectral projector of B on its eigen space associated to the eigen value 0, which is $\mathcal{N}(B)$ because B is a generator of a Markov chain.(see prop. 5 ch. 6 Pallu de la Barrière[15]);

$R\mu(A)$ the resolvent of the operator A is by definition $(A-\mu)^{-1}$ where $\mu \in \mathbb{C}$; $A\mu = A - \mu I$; $C = R\mu(A)B$.

II.1. Analytical Result

<u>Theorem 1</u>

<u>V_ε solution of (1.1) admits the expansion</u> $\sum_{i=0}^{+\infty} \varepsilon^i V_i(x)$ <u>with</u>

$V_i = \tilde{V}_i + \overline{V}_i$ where $\tilde{V}_i \in \mathcal{R}(B)$ and $\overline{V}_i \in \mathcal{N}(B)$.

The sequence $(\tilde{V}_i, \overline{V}_i)$ is uniquely determined by :

$$(1.4) \qquad B\tilde{V}_i + A\mu V_{i-1} + f_{i-1} = 0 \ , \ \tilde{V}_i \in \mathcal{R}(B),\ i = 1, 2, \ldots, \tilde{V}_0 = 0 \ ;$$

$$(1.5) \qquad PA\mu\overline{V}_i + PA\mu\tilde{V}_i + Pf_i = 0 \ , \ \overline{V}_i \in \mathcal{N}(B),\ i = 0, 1, \ldots, \ .$$

Before proving this theorem let us give a

lemma 1

The operator $R\mu(A)B$ has the eigen value 0 and the nilpotent operator associated is zero

Proof of the lemma. B being a generator $R\mu(A)B1 = 0$ and 0 is eigen value of $R\mu(A)B$. Given $g : \mathbb{E} \to \mathbb{R}$, let us consider

$$W_\varepsilon = \mathbb{E}\sum_{t=0}^{+\infty} \frac{\varepsilon}{(1+\mu\varepsilon)^{t+1}} \ g \circ X_t$$

W_ε is bounded by $\frac{1}{\mu} \sup_{x\in E} |g(x)|$. This bound is independent of ε . But W_ε is solution of the Kolmogorov equation :

$$A\mu W_\varepsilon + \frac{1}{\varepsilon} BW_\varepsilon + g = 0$$

and

$W_\varepsilon = - \varepsilon R_{-\varepsilon}(C)\ R\mu(A)g$ by definition of the resolvent of an operator. By Kato [11] ch. I.5.3. we obtain that W_ε is bounded if and only if the nilpotent associated to the eigen value 0 of the operator C is zero, and the result is proved.

Proof of the Theorem. The solution of (1.3) can be written

$$V_\varepsilon = - \varepsilon R_{-\varepsilon}(C)\ R\mu(A)f_\varepsilon$$

Using Kato [11] chap. I.5.3. we know that $-\varepsilon R_{-\varepsilon}(C)$ is analytic and its convergence radius r is the smallest modulus of the non zero eigen-value of C, and the convergence radius of V_ε is r, because $\sup_{x\in E,i} f_i(x) \le C_f$.

Let us prove now that $(\tilde{V}_i, \bar{V}_i)$ solution of (1.4), (1.5), is uniquely determined.

Let us prove this recursively for that let us first verify that

$$(1.6) \qquad PA\mu P_C R\mu(A) = P$$

where P_C denoted the spectral projector on the 0-eigen-space of C. But by lemma 1 we have $\mathcal{N}(c) \cap \mathcal{R}(c) = 0$ which implies that $A\mu\mathcal{N}(B) \cap \mathcal{R}(B) = 0$ because $A\mu$ is regular.

Now let us take $f \in \mathcal{R}(B)$ then it exists g such that $Bg = f$ and we have

$$(1.7) \qquad A\mu P_C R\mu(A)f = A\mu P_C R\mu(A)Bg = A\mu P_C Cg = 0$$

If we take $f \in A\mu\,\mathcal{N}(B)$, $\exists\, g \in \mathcal{N}(B)$: $f = A\mu g$ and

$$(1.8) \qquad A\mu P_C R\mu(A)f = A\mu P_C R\mu A\mu g = A\mu P_C g = A\mu g = f$$

(1.7), (1.8) imply that $A\mu P_C R\mu(A)$ is the projection on $A\mu\mathcal{N}(B)$ along $\mathcal{R}(B)$ which implies (1.6).

(1.6) shows that $PA\mu$ is invertible in $\mathcal{N}(B)$ and its inverse is $P_C R\mu(A)$ so $\bar{V}_i = - P_C(\tilde{V}_i + R\mu(A)\ f_i) = - P_C R\mu(A)(A\mu\tilde{V}_i + f_i)$ is solution of (1.5).

(1.5) can be written by definition of $V_i = \bar{V}_i + \tilde{V}_i$: $P(A_\mu V_i + f_i) = 0$, this relation implies that $A\mu V_i + f_i \in \mathcal{R}(B)$ which proves that there exists a solution to $(1.4)_{i+1}$.

So the sequence $\{V_i\}$ is well defined by induction.

Now the convergence radius of the series $\sum_{i=0}^{+\infty} \varepsilon^i V_i$ where V_i is defined by $(1.4)_i$, $(1.5)_i$, is r defined at the begining of the proof. This can be proved recursively writting $(1.4)_i$.

$$R\mu(A)B\,\tilde{V}_i + V_{i-1} + R_\mu(A)f_{i-1} = 0$$

and because $\sup_{i,x\in E} f_i(x) \leq C_f$

Now $V_\varepsilon = \sum_{i=0}^{+\infty} \varepsilon^i V_i$ where V_i are defined by $(1.4)_i$ $(1.5)_i$ satisfies

$\frac{1}{\varepsilon} B\, V_\varepsilon + A\mu V_\varepsilon + f_\varepsilon = 0$ because $(1.4)_i$ and $(1.5)_i$ implies that $BV_i + A\mu V_{i-1} + f_{i-1} = 0$ and $BV_0 = 0$, then the proof is achieved.

II.2. Probabilistic interpretation of the terms of the expansion of V_ε

For that let us define an aggregated markov chain denoted by $\bar{X}_t$. Its states are the final classes of the fast chain denoted by $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ and we call $\bar{E}$ the set of the final classes of the fast chain $\bar{E}=\{\bar{x}_1, \ldots, \bar{x}_m\}$. Its generator is defined by $\bar{A}$ with

$$(1.8) \qquad \bar{a}_{\bar{x}\bar{x}'} = \sum_{x\in\bar{x}} p_{\bar{x}}(x) \; \{\sum_{x'\in\bar{x}'} \bar{a}_{\bar{x}\bar{x}'} + \sum_{y\in\bar{y}} q_{\bar{x}'}(y)\; a_{xy}\}$$

where : $p_{\bar{x}} : x \to \mathbb{R}^+$ is the invariant measure of the fast chain of support $\bar{x}$,

$\bar{y}$ is the set of transient states of the fast chain,

$q_{\bar{x}'}(y) \in \mathbb{R}^+$ is the probability to end in the final class $\bar{x}'$ starting from the transient state y for the fast chain.

We define an aggregated cost :

$$(1.9) \qquad \bar{g} : \bar{E} \to \mathbb{R}$$

$$\bar{x} \to \sum_{x\in\bar{x}} g(x)\; p_{\bar{x}}(x)$$

Then we have the :

Theorem 2

$\bar{V}_i$ and $\tilde{V}_i$ defined by $(1.4)_i$ and $(1.5)_i$ have the following probabilistic interpretations

$$(1.10) \qquad \bar{V}_i(x) = \bar{V}_i(\bar{x}) = \mathbb{E}^{\bar{x}}_{ag} \sum_{t=0}^{+\infty} \frac{1}{(1+\mu)^{t+1}} \bar{g}_i \circ \bar{X}_t, \forall x \in \bar{x}, \; \forall \bar{x} \in \bar{E}, \; \forall i \in \mathbb{N},$$

$$(1.11) \qquad \bar{V}_i(y) = \sum_x q_{\bar{x}}(y) \bar{V}_i(x) = E^y_{fast} V(Z_\tau), \; \forall y \in \bar{y}, \quad \forall i \in \mathbb{N} ;$$

$$(1.12) \qquad \tilde{V}_i(z) = \mathbb{E}_{fast} \lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} (T - t) \, h_i \circ Z_t \qquad \forall i \in 1, 2, \ldots ;$$

with $\tau = \inf \{t \geq 0, Z_t \in \bar{E}\}$, where ag means "for the aggregated chain", fast "for the fast chain" and Z_t is the fast chain and :

$$(1.13) \qquad g_i = A\mu\tilde{V}_i + f_i ;$$

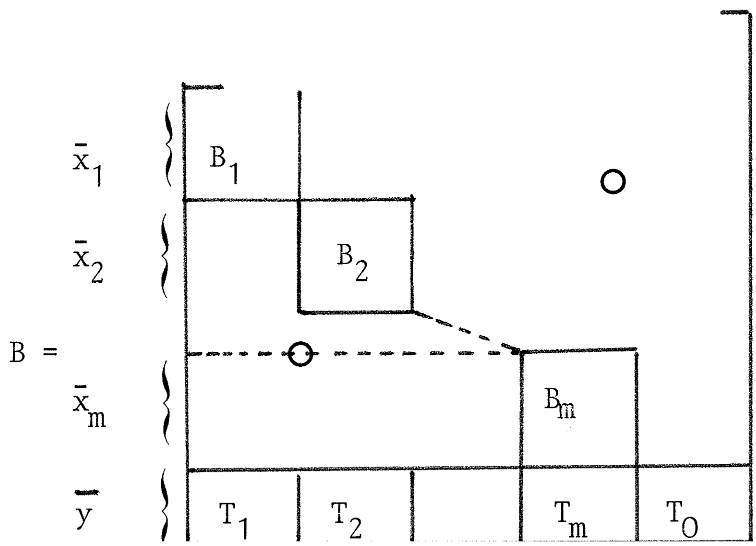
$$(1.14) \qquad h_i = A\mu V_{i-1} + f_{i-1} .$$

Proof :

Using the definition (1.14) of h_i $(1.4)_i$ can be written $B\tilde{V}_i + h_i = 0$, B being the generator of the fast chain the theorem 5.14 of Kemeny Snell [12] extended to the general situation th.4 ch.6 Pallu de la Barriere [15] gives the interpretation (1.12).

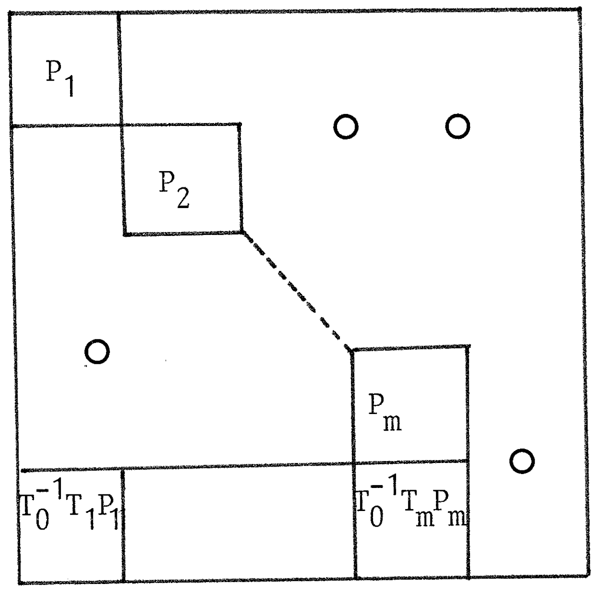
Interpretation of $\bar{V}_i$.

B can be written using the partition of $E = \bar{x}_1 \cup \bar{x}_2 \cup \ldots \cup \bar{x}_m \cup \bar{y}$



P is given fig. 3 p. 46 Lanery [13] or Kemeny Snell [12]

P =



With $P_i = \mathbb{1}_{\bar{x}_i} \otimes p_{\bar{x}_i}$

The interpretation of $T_0^{-1}\ T_i\ 1_{\bar{x}_i}(y)$ is the probability $q_{\bar{x}}(y)$ to end in the final class $\bar{x}$, starting from the transient state $y \in \bar{y}$ for the fast chain. It follows that

$$T_0^{-1}\ T_i\ P_i = q_{\bar{x}_i} \otimes p_{\bar{x}_i}$$

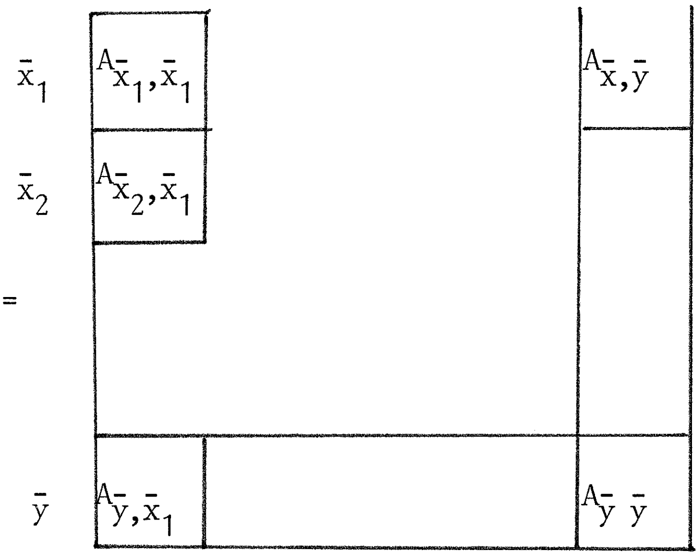
Now $P\bar{V}_i = \bar{V}_i$ because $\bar{V}_i$ belongs to $\mathcal{N}(B)$ so

$PA\mu P\bar{V}_i = PA\mu\bar{V}_i$, but $PA\mu P = PAP - \mu P = \bar{A} - \mu P$

with $\bar{A}$ the following generator.

$\bar{A}_{\bar{x}\bar{x}'} = \mathbb{1}_{\bar{x}} \otimes p_{\bar{x}'}\ (p_{\bar{x}}\ (A_{\bar{x}\bar{x}'} \mathbb{1}_{\bar{x}'} + A_{\bar{x}\bar{y}} q_{\bar{x}'})) = \mathbb{1}_{\bar{x}} \otimes p_{\bar{x}'} \bar{a}_{\bar{x}\bar{x}'}$ where $\bar{a}_{\bar{x}\bar{x}'}$ is defined by (1.8) using the notation

A =



$\bar{A}_{\overline{xy}} = 0$ so the solution of $(1.6)_i$ needs only the knowledge of the restriction of $\bar{A}$ to $\bar{E}$, indeed because

$\bar{A}_{y\bar{x}'} = \sum_{\bar{x}} q_{\bar{x}}(y)\ \bar{a}_{\overline{xx}'}$, $y \in \bar{y}$, we see that the rows of $\bar{A}$ for $y \in \bar{y}$ are linear combinations of the other ones.

$\bar{A}$ restricted to $\bar{E}$ denoted by $A_{\bar{E}}$ is the generator of a Markov chain which can be lumped using th. 6.3.2. of Kemeny Snell [12] . This lumped chain admits m states $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m)$ and its transition probabilities are given by (1.8). We still denote by $\bar{A}$ the generator of the lumped chain.

We can check that $\bar{g}(\bar{x})$ defined by (1.9) is equal to $Pg(x)$, $\forall x \in \bar{x}$, $\forall \bar{x} \in \bar{E}$

We can verify using the relation-ship $Pg(y) = \sum_{\bar{x}} q_{\bar{x}}(y)\bar{g}(\bar{x})$, $\forall\, y \in \bar{y}$, the compatibility of the system $(1.5)_i$.

The condition $P\bar{V}_i = \bar{V}_i$ determines the values of $\bar{V}_i(y)$, $\forall\, y \in \bar{y}$:

$$\bar{V}_i(y) = \sum_{\bar{x} \in \bar{E}} q_{\bar{x}}(y)\ \bar{V}_i(\bar{x}).$$

The interpretation (1.11) follows from the relation :

$$\mathbb{E}^{y}_{fast}\ \bar{V}_i(Z_\nu) = \sum_{\bar{x}} \bar{V}_i(\bar{x})\ \mathbb{P}_{fast}\ (Z_\nu \in \bar{x}) = \sum_{\bar{x}} \bar{V}_i(\bar{x})\ q_{\bar{x}}(y)$$

Remark 1 $(1.5)_i$ can be written

(1.15) $\quad (A\mu\bar{V}_i + A\mu\tilde{V}_i + f_i,\ p) = 0\ \forall\ p \in \eta\ (B^*),\quad \bar{V}_i \in \eta\ (B).$

Taking the extremal invariant probabilities of the fast chain as a basis of $\eta\ (B^*)$, and $(q_{\bar{x}},\ \bar{x} \in E)$ the probabilities to end in the final class for the fast chain as a basis of η (B) we obtain the interpretation of $\bar{V}_i$ (th 2 Delebecque-Quadrat [7]) as the solution of a Kolmogorov equation for the aggregated chain.

Remark 2 The classical small discount problem :

$$V_\varepsilon(x) = \mathbb{E}^x \sum_{t=0}^{\infty} \frac{\varepsilon}{(1+\varepsilon)^{t+1}}\ f(X_t)$$

where X_t is a markov chain without weak and strong interactions can be seen like a particular case A = 0 indeed V_ε in this case is solution of :

$$\left(- I + \frac{B}{\varepsilon} \right) V_\varepsilon + f = 0$$

where B is the generator of the Markov chain X_t.

II.3 Computational aspects

Using the equation $(1.4)_i$ and $(1.5)_i$ the sequence V_i can be computed by a two level algorithm :

level 1 (fast)

- Solve m decoupled systems of $n_{\bar{x}}$ linear equations, $n_{\bar{x}}$ is the number of states in $\bar{x}$.

Each system is define

$B_{\bar{x}} \tilde{V}_i^{\bar{x}} + h_i^{\bar{x}} = 0$, $p_{\bar{x}} \tilde{V}_i^{\bar{x}} = 0$, where $B_{\bar{x}}$ [resp $\tilde{V}_i^{\bar{x}}$, $h_i^{\bar{x}}$] is the restriction of B [resp $\tilde{V}_i$, h_i] to the final class $\bar{x}$.

- Solve the following system of equations on the transient classes of the fast chain :

$$\begin{cases} B_{\bar{y}} \tilde{V}_i + h_i = 0, \quad \forall y \in \bar{y} \; ; & * \\ \tilde{V}_i(x) = \tilde{V}_i^{\bar{x}}(x), \; \forall x \in \bar{x}, \; \forall \bar{x} \in \bar{E}. \end{cases}$$

The solution gives the values of $\tilde{V}_i$ on the transient states the computational cost being the solving of a set of $n_{\bar{y}}$ linear equations, where $n_{\bar{y}}$ is the number of transient states of the fast chain.

- Compute the invariant prob. measures of support the final classes, that is solve $p_{\bar{x}} B_{\bar{x}} = 0 \quad (p_{\bar{x}}, 1) = 1 \quad \forall \bar{x} \in \bar{E}$.

- Compute the probability to end in a final class starting from a transient states that is solve :

$$\left\{ \begin{array}{ll} B_{\bar{y}} q_{\bar{x}} = 0 \text{ on } y \in \bar{y} & * \\ & , \; \forall \bar{x} \in \bar{E}. \\ q_{\bar{x}} = 1_{\bar{x}} \quad \text{on } \bar{E} & \end{array} \right.$$

Level 2

- Define the aggregated chain of generator given by the formula (1.8).

- Solve $\bar{A}_\mu \bar{V}_i + \bar{g}_i = 0$ which is a system of m equations.

- Define $\bar{V}_i$ on the transient states by the formula $\bar{V}_i(y) = \sum_{\bar{x}} q_{\bar{x}}(y) \bar{V}_i(\bar{x})$

---

* $B_{\bar{y}}$ is the rectangular matrix obtained by taking the lines of B, the index of which to $\bar{y}$.

Remark 4

The size of the largest linear system that we have to solve is

$$\sup \left(\sup_{\bar{x} \in \bar{E}} n_{\bar{x}}, n_{\bar{y}}, m \right)$$

If the inversion is done by a Gauss method and $n_y \simeq n_x \simeq m \simeq \sqrt{n}$ then the cost is of order $k \sqrt{n} (\sqrt{n})^3 \simeq kn^2$. So we save much computation time by this method for a large scale system.

III. OPTIMAL CONTROL OF MARKOV CHAIN WITH STRONG AND WEAK INTERACTIONS

III.1 Analytical Result

Given :

- $t \in N$ the time ;
- U a finite number set called the set of controls, $S : E \to U$ is called a strategy ; and $\mathcal{S} = \{S : E \to U\}$ is the set of strategies ;
- X_t a controlled markov chain, that is its transition matrix is a function of $u \in U$, $M(u)$;

  We denote by M^S the matrix $M^S_{xx'} = M_{xx'}(S(x))$, we denote by X^S_t the markov chain of transition matrix M^S, we suppose that the markov chain has weak and strong interactions (that is : there exists $\varepsilon > 0$ such that $M(u) - I = B(u) + \varepsilon A(u)$, with $B(u)$ and $A(u)$ being generators of markov chains $B(u) + I$ and $A(u) + I$ are transition matrices), we use also the notation A^S [resp B^S] for the matrix $A^S_{xx'} = A_{xx'}(S(x))$ [resp $B^S_{xx'} = B_{xx'}(S(x))$] ;

- $f : \underset{i}{\mathbb{N}} \times \underset{x}{E} \times \underset{u}{U} \to \mathbb{R}^+$ with $\sup_{i,x,u} f(i, x, u) \le C_f$

For the strategy S, $\varepsilon \in R^+$, let us denote by f^S_ε the function :

$$f^S_\varepsilon : \mathbb{E} \to \mathbb{R}^+ \qquad x \to f^S_\varepsilon(x) = \sum_{i=0}^{+\infty} \varepsilon^i f_i(s, S(x))$$

which is called the cost function,

and f^S the function $f^S : \underset{i}{\mathbb{N}} \times \underset{x}{E} \to \mathbb{R} \quad f_i(x, S(x))$.

We study the stochastic control problem for $\mu \in R^+$, $\mu > 0$:

$$\underset{S \in \mathcal{S}}{\mathrm{Min}}\ \mathbb{E} \sum_{t=0}^{+\infty} \frac{\varepsilon}{(1+\varepsilon_\mu)^{t+1}} f^S_\varepsilon \circ X^S_t$$

V^S_ε denotes the function :

$$V^S_\varepsilon \quad \mathbb{E} \to \mathbb{R}^+$$
$$x \to \mathbb{E}\{ \sum_{t=0}^{+\infty} \frac{\varepsilon}{(1+\varepsilon\mu)^{t+1}} \; f^S_\varepsilon \circ X^S_t \mid X^S_0 = x\} \quad ,$$

and $V_\varepsilon = \underset{S\in \mathcal{S}}{\text{Min}} \; V^S_\varepsilon)$ (Componentwise).

The purpose of this chapter is to give the expansion in ε of

$$V_\varepsilon = \sum_{i=0}^{+\infty} \varepsilon^i V_i .$$

We know by chapter 1 that V^S_ε has the expansion $V^S_\varepsilon = \sum_{i=0}^{+\infty} \varepsilon^i V^S_i$.

$V^{S,\ell}$ [resp $V^{S,\ell}_\varepsilon$, resp V^ℓ , resp V^ℓ_ε] denotes the sequence $(V^S_0, V^S_1, \dots, V^S_\ell)$ [resp $\sum_{i=0}^{\ell} \varepsilon^i V^S_i$, resp $(V_0, V_1, \dots, V_\ell)$, resp $\sum_{i=0}^{\ell} \varepsilon^i V_i$] .

P^S is the spectral projector on the 0-eigen space of B^S.

Let us note by $\succcurlyeq$ the lexicographic order defined on a finite or infinite sequence of numbers, the minimum for this order relation will be denoted by $\overrightarrow{\text{Min}}$.

For two given strategies S, S', let us define the functions :

$$H^S_0 : \mathbb{R}^{E\times N} \to \mathbb{R}^E$$
$$y = (y_0, y_1, \dots) \mapsto B^S y_0 = H^S_0(y) \quad ;$$

$$H^S_i : \mathbb{R}^{E\times N} \to \mathbb{R}^E$$
$$y = (y_0, y_1, \dots) \mapsto A^S_\mu y_{i-1} + B^S y_i + f^S_{i-1} = H^S_i(y), \; i \in N - \{0\} \; ;$$

$$H_i^{SS'} : \mathbb{R}^{E\times N} \to \mathbb{R}^E \qquad ; \qquad y \quad H_i^S(y) - H_i^{S'}(y)$$

We shall use also the following notations :

$$H^S = (H_i^S,\ i \in \mathbb{N}) \ ; \ H^{SS'} = (H_i^{SS'},\ i \in \mathbb{N}) \ ;$$

$$H^{S,\ell} = (H_i^S,\ i = 0, \ldots, \ell) \ ; \ H_i^{SS',\ell} = (H^{SS'},\ i = 0, \ldots, \ell) \ ;$$

$$H_\varepsilon^S = \sum_{i=0}^{+\infty} \varepsilon^i H_i^S \ ; \ H_\varepsilon^{SS'} \sum_{i=0}^{+\infty} \varepsilon^i H_i^{SS'} \ ;$$

$$H_\varepsilon^{S,\ell} = \sum_{i=0}^{\ell} \varepsilon^i H_i^S \ ; \ H_\varepsilon^{SS',\ell} = \sum_{i=0}^{\ell} \varepsilon^i H_i^{SS'} \ .$$

We have the :

Lemma 2

$\forall x \in E$, $V^S(x) \succcurlyeq 0$ [resp $V^{S,\ell} \succcurlyeq 0$] $\Leftrightarrow \exists \delta : \forall \varepsilon \le \delta,\ \ \varepsilon \ge 0,\ V_\varepsilon^S(x) \ge 0\ \ \forall x \in E$ [resp $V_\varepsilon^{S,\ell}(x) \ge 0$] .

Proof

The necessity being trivial let us prove the sufficiency of the condition.

It is sufficient to prove that :

$$(k \in \mathbb{N}\ \ V^{S,k}(x) = 0 \Rightarrow V_{k+1}^S(x) \ge 0).$$

But $V^{S,k}(x) = 0 \Rightarrow V_\varepsilon^{S,k}(x) = 0 \Rightarrow \lim\limits_{\substack{\varepsilon\to 0\\ \delta\ge\varepsilon>0}} \frac{1}{\varepsilon^{k+1}} V_\varepsilon^S(x) = V_{k+1}^S(x)$ and because $\varepsilon > 0$ and $V_\varepsilon^S(x) \ge 0$ we have $V_{k+1}^S(x) \ge 0$.

The following result is a generalized Howard [10], Miller-Veinott [14] algorithm for the situation where we have strong and weak interactions.

Theorem 3 $\exists\delta \quad \forall\varepsilon \leq \delta$, $\varepsilon \geq 0$ we have :

1) $H^{S,S'} \circ V^S(x) \succcurlyeq 0 \quad \forall x \in E \Rightarrow V^S(x) \succcurlyeq V^{S'}(x) \quad \forall x \in E \Leftrightarrow V^S_\varepsilon(x) \geq V^{S'}_\varepsilon(x) \, \forall x \in E$

2) $\ell \geq 1$, $H^{SS',\ell} \circ V^S(x) \succcurlyeq 0, \forall x \in E \Rightarrow V^{S,\ell-1}(x) \succcurlyeq V^{S',\ell-1}(x) \quad \forall x \in E \Leftrightarrow$

$V^{S,\ell-1}_\varepsilon(x) \geq V^{S',\ell-1}_\varepsilon(x) \quad \forall x \in E$.

3) V_ε admits an expansion $\sum_{i=0}^{+\infty} \varepsilon^i V_i$, $\quad V = (V_0, V_1, \ldots, V_n, \ldots)$ satisfies the vector Hamilton Jacobi equation : $\vec{\mathrm{Min}}_{S \in \mathcal{S}} H^S \circ V(x) = 0, \forall x \in E$.

4) The vector $V^{\ell-1} = (V_0, V_1, \ldots, V_{\ell-1})$ is uniquely determined by the equation $\vec{\mathrm{Min}}_{S \in \mathcal{S}} \; H^{S,\ell} \circ V(x) = 0$.

Proof

It is a straightforward adaptation of the techniques used by Miller-Veinott [14].

V^S_ε satisfies the Kolmogorov equation :

(2.1) $$0 = B^S V^S + \varepsilon A^S V^S + \varepsilon f^S_\varepsilon \, .$$

We have also

(2.2) $$0 = B^{S'} V^{S'} + \varepsilon A^{S'} V^{S'} + \varepsilon f^{S'}_\varepsilon$$

using the expansion of V^S_ε and $V^{S'}_\varepsilon$ which exist we have

$$0 = \sum_{i=0}^{+\infty} \varepsilon^i (H^S_i(V^S) - H^{S'}_i(V^{S'})) \Rightarrow$$

(2.3) $$0 = \sum_{i=0}^{+\infty} \varepsilon^i \; [H^S_i(V^S) - H^{S'}_i(V^S) + H^{S'}_i(V^S) - H^{S'}_i(V^{S'})]$$

and denoting by

$$W_\varepsilon^{SS'} = \sum_{i=0}^{+\infty} \varepsilon^i (V_i^S - V_i^{S'}) \text{ and } W^{SS'} = (W_0^{SS'}, W_1^{SS'}, \ldots, W_n^{SS'}, \ldots)$$

(2.3) can be written $0 = H_\varepsilon^{SS'}(V^S) + B^{S'} W_\varepsilon^{SS'} + \varepsilon A_\mu W_\varepsilon^{SS'}$

So we have the stochastic interpretation

(2.4) $$W_\varepsilon^{SS'} = E \sum_{n=0}^{+\infty} H_\varepsilon^{SS'} \circ V^S \circ X_t^{S'}$$

Now let us prove the result 1)

$H^{SS'} \circ V^S(x) \succcurlyeq 0 \; \forall x \in E \Rightarrow \exists \delta \; \forall \varepsilon \leq \delta, \varepsilon \geq 0 \; H_\varepsilon^{SS'} \circ V^S(x) \geq 0 \quad \forall x \in E$ by lemma 2 $\Rightarrow W_\varepsilon^{SS'}(x) \geq 0, \forall x \in E, \forall \varepsilon \leq \delta$, by the maximum principle or the interpretation 2.4 $\Rightarrow W^{SS'}(x) \succcurlyeq 0 \;\; \forall x \in E$ by the lemma 2 so the 1) is proved.

Let us prove the part 2)

$\ell \geq 1 \;\; H^{SS',\ell} \circ V^S(x) \succcurlyeq 0 \; \forall x \in E \;\; \Rightarrow \exists \delta, k : \forall \varepsilon \leq \delta, \varepsilon \geq 0 \;\; H_\varepsilon^{SS'} \circ V^S(x) + k\varepsilon^{\ell+1} \geq 0 \;\; \forall x \in E$ by lemma 2 $\Rightarrow$ by interpretation (2.4) and majoration $W_\varepsilon^{SS'}(x) + k\varepsilon^\ell \geq 0 \quad \forall x \in E \Leftrightarrow V^{S,\ell-1}(x) \succcurlyeq V^{S',\ell-1}(x), \; \forall x \in E$ by lemma 2 $\Leftrightarrow V_\varepsilon^{S,\ell-1}(x) \geq V_\varepsilon^{S',\ell-1}(x), \; \forall x \in E, \;\; \forall \varepsilon \in \delta$.

For the part 3),

first let us prove the existence of a solution of $\overrightarrow{\operatorname{Min}}_{S \in \mathcal{S}} H^S \circ V(x) = 0$.

For that we use the strategy iteration algorithm :

$$S \xrightarrow{\text{solve } H^S \circ V^S = 0} V^S \xrightarrow[S' \in \mathcal{S}]{\text{solve Min } H^{S'} \circ V^S} S'$$

by this way we obtain a sequence of strategies S_n and a sequence of V^{S_n}.

By the first part we have

$$\forall x \in E \quad V^{S_n}(x) \leqslant V^{S_{n-1}}(x) \text{ because } H^{S_{n-1}S_n} \circ V^{S_{n-1}} \leqslant 0.$$

So after a finite number of iterations $V^{S_n} = V^{S_{n-1}}$ because the number of strategies is finite. We denote $S^* = S^n$ then

$$(2.5) \qquad H^{S^*} \circ V^{S^*}(x) = 0 \quad \forall x \in E$$

and

$$(2.6) \qquad H^{SS^*} \circ V^{S^*}(x) \geqslant 0 \quad \forall x \in E, \forall S \in \mathcal{S}$$

and (2.5), (2.6) $\Leftrightarrow$

$$\overrightarrow{\text{Min}}_{S \in \mathcal{S}} \quad H^S \circ V^{S^*}(x) = 0, \forall x \in E.$$

Now because of (2.6) and part 1) we have $V^{S^*}(x) \leqslant V^S(x), \forall x \in E, \forall S \in \mathcal{S}$ and lemma 1 $\Rightarrow \exists \delta$, $V_\varepsilon^{S^*}(x) \leq V_\varepsilon^S(x), \forall x \in E, \forall \varepsilon \leq \delta, \quad \varepsilon > 0$. So $V_\varepsilon^{S^*} = V_\varepsilon$.

The solution V of $\overrightarrow{\text{Min}}_{S \in \mathcal{S}} H^S \circ V$ is unique because by part 1 if there are two solutions V^1 and V^2 we would have $V^1 \geqslant V^2$ and $V^2 \geqslant V^1 \Rightarrow V^1 = V^2$ by antisymmetry of the order relation .

Let us prove now the part 4)

We prove first the existence and uniqueness of

$$\overrightarrow{\text{Min}}_{S \in \mathcal{S}} \quad H^{S,\ell} \quad V(x) = 0 \qquad \forall x \in E$$

by the same technique using this time part 2.

By part 2 this solution V^{S_n} satisfies

$$V^{S_n,\ell-1}(x) \leqslant V^{S,\ell-1}(x) \quad \forall x \in E, \quad \forall S \in \mathcal{S} \quad .$$

So $V_\varepsilon^{S_n,\ell-1}(x) \leqslant V_\varepsilon^{S,\ell-1}(x)$, $\forall x \in E$, $\forall S \in \mathcal{S}$, and this implies that

$$V_\varepsilon^{S_n,\ell-1}(x) = \sum_{i=0}^{\ell-1} \varepsilon^i V_i(x)$$ where V_i are the terms of the expansion of V_ε

III.2 Probabilistic interpretation of the expansion of V_ε

If V is the expansion of V_ε , we denote by $\mathcal{S}_i = \underset{S\in \mathcal{S}_{i-1}}{\mathrm{Argmin}}\, \vec{H}_i^S \circ V, \mathcal{S}_0 = \mathcal{S}$.

For each strategy we introduce the fast chain Z_t^S of generator B^S and the aggretated chain $\bar{X}_t^S$ of generator $P^S A^S P^S$, we denote by $\bar{x}^S$ the corresponding aggregated state, and $q^y_{\bar{x}^S}$ the probability to finish in $\bar{x}^S$ starting from the fast-transient state y. $\bar{y}^S$ is the set of fast-transient states.

We have the :

Theorem 4

If we denote by

(2.7) $\quad h_i^S = A_\mu^S V_{i-1} + f_{i-1}^S \quad , i = \mathbb{N}-\{0\} \quad , \quad h_0^S = 0$;

(2.8) $\quad \bar{g}_i^S = P^S A_\mu^S \tilde{V}_i + P^S f_i^S \quad , \quad i \in \mathbb{N}$;

We have for $i \in \mathbb{N}$

(2.9) $\quad \bar{V}_i^S(x) = \bar{V}_i^S(\bar{x}^S) = P^S V_i(x) \leq E_{Ag}^{\bar{x}^S} \sum_{t=0}^{\infty} \frac{1}{(1+\mu)^{t+1}} \bar{g}_i^S o \bar{X}_t^S, \ \forall\, S\in\mathcal{S}_{i+1}\ \forall x\in \bar{x}^S$;

$\quad \bar{V}_i^S(y) = \sum_{\bar{x}^S} q^y_{\bar{x}^S}\ \bar{V}_i^S(\bar{x}^S), \ \forall\ y \in \bar{y}^S$;

(2.10) $\quad \tilde{V}_i^S = V_i - \bar{V}_i^S \leq \mathbb{E}_{fast} \lim_{T\to\infty} \sum_{t=0}^{T-1} (1-\frac{t}{T})\ h_i^S \circ Z_t, \ \forall S \in \mathcal{S}_i$.

Corollary

In the particular case where $\eta(B^S)$ is independent of the strategy $V_0^S = V_0$ is independent of S and (2.9) becomes :

$$(2.11) \qquad V_0(x) = \bar{V}_0(\bar{x}) = \underset{S \in \mathcal{S}_0}{\mathrm{Min}} \; \mathbb{E}_{Ag} \sum_{t=0}^{+\infty} \frac{1}{(1+\mu)^{t+1}} \bar{f}_o^S \circ \bar{X}_t^S \, , \quad \forall x \in \bar{x} \; ;$$

$$V_0(y) = \sum_{x} q_x^y \, V_i(\bar{x}) .$$

Remark : (2.11) can be cut in a long run control problem an a short run one (see Delebecque-Quadrat [7]), using a decomposition by the quantities.

Proof

Let us take $S \in \mathcal{S}_{i+2}$, $S' \in \mathcal{S}_{i+1}$, V the expansion of V_ε. By theorem 3 we have :

$$(2.13) \qquad 0 = A_\mu^S V_i + B^S V_{i+1} + f_i^S \le A_\mu^{S'} V_i + B^{S'} V_{i+1} + f_i^{S'}$$

multipliying (2.13) by $P^{S'}$ we obtain :

$$(2.14) \qquad 0 \le P^{S'} A_\mu^{S'} V_i + P^{S'} f_i^{S'}$$

with $P^{S'} V_i = \bar{V}_i^{S'}$ and $\tilde{V}_i^{S'} = V_i^{S'} - \bar{V}_i^{S'}$ (2.14) can be written

$$0 \le P^{S'} A_\mu^{S'} P^{S'} \bar{V}^{S'} + P^{S'} f_i^{S'} + P^{S'} A_\mu^{S'} \tilde{V}^{S'} .$$

Using theorem 2 $P^{S'} A^{S'} P^{S'}$ can be interpreted as the generator of an lumpable chain. We have :

$$\bar{V}^{S'}(\bar{x}^{S'}) \le \mathbb{E}_{Ag} \sum_{t=0}^{+\infty} \frac{1}{(1+\mu)^{t+1}} \bar{g}_i^{S'} \circ \bar{X}_t^{S'} \; ;$$

$$\bar{V}^{S'}(y) = \sum_{\bar{x}^{S'}} q_{\bar{x}^{S'}}^y \, \bar{V}^{S'}(\bar{x}^{S'}) \, , \quad \forall y \in \bar{y}^{S'} .$$

Now we have for $S \in \mathcal{S}_i$

$$(2.15) \qquad \mathbb{E}_{fast}(V_i o Z_T^S - V_i o Z_0^S) = \mathbb{E}_{fast}^{Z_0^S} \sum_{t=0}^{T-1} B^S V_i o Z_t^S \geq$$

$$\geq - \mathbb{E}_{fast}^{Z_0^S} \sum_{t=0}^{T-1} A_\mu^S V_{i-1} o Z_t^S + f_{i-1}^S o Z_t^S \quad \text{because } (2.13)_{i-1}$$

Summing (2.15) for $T = 1, \ldots, N$ we obtain

$$\mathbb{E}_{fast} \frac{1}{N} \sum_{t=1}^{N} (V_i o Z_t^S - V_i o Z_0^S) \geq - \mathbb{E}_{fast}^{Z_0^S} \sum_{t=0}^{N-1} (1 - \frac{t}{N}) h_i^S o Z_t^S$$

$$\text{But } \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} V_i o Z_t^S = P^S V_i$$

this implies that :

$$\tilde{V}_i^S = V_i - P^S V_i \leq \mathbb{E}_{fast}^{Z_0^S} \lim_{T \to \infty} \sum_{t=0}^{T-1} (1 - \frac{t}{T}) h_i^S o Z_t^S$$

Proof of the corollary

Suppose now that $\eta(B^S)$ is independent of the strategy S it follows that $\bar{x}^S$ is independant of S, but $\tilde{V}_0^S = 0$ which implies that $\bar{V}_0^S = V_0$ and the result follows. using the fact that on $\bar{E}$ $\mathcal{S}_0 = \mathcal{S}_1$.

Remark

The result of Veinott [18] is the particular case A = 0. The case of U infinite set could be handled by techniques used in CHITASHVILLI [3] but by this way we shall not have the complete expansion of V_ε. (and some proofs are not completely clear). Another particular case with U infinite set can be found in DELEBECQUE-QUADRAT [7].

III.3 Computational aspects

The theorem 3 can be used to obtain an algorithm giving the first ℓ terms of the expansion of V_ε, by solving :

1) given a strategy S compute V^S, this step can be done by coordination-decomposition algorithm described in I_3) after determining the final class of B^S.

2) $\vec{\text{Min}}_{S' \in \mathcal{S}} \; H^{S',\ell}(V^S)$, this is a local minimisation.

3) go to 1) until convergence occurs.

The largest computation that we have to do is the computation of the invariant measure of the largest final class of the fast chain, or the computation of the aggregated cost V^S, but we never have to solve a problem of the size of the initial problem (if there are several final classes of the fast chain). In the classical problem A = 0 we don't save computation time but we solve a difficult problem. In the case considered here we can hope to save computation time which was in fact the purpose of this study.

REFERENCES

[1] A. BENSOUSSAN-J.L. LIONS-G. PAPANICOLAOU
"Asymptotic Analysis for Periodic Structures"
North Holland, 1978.

[2] R. BELLMAN
"A Markovian Decision Process"
Jour. Math. Mech. (1957), 6, p. 679-684.

[3] R.Y. CHITASHVILLI
"A Controlled Finite Markov Chain with an Arbitrary Set of Decisions"
Siam, th. of Probability 20, 4, (1975), p. 839-846.

[4] J.H. CHOW-P.F. KOKOTOVIC
"Decomposition of Near Optimal State Regulator for System with Slow and Fast Modes"
IEEE A.C. 1976.

[5] P.J. COURTOIS
" Decomposability "
ACM Monograph series, Academic Press 1977.

[6] F. DELEBECQUE-J.P. QUADRAT
"Contribution of Stochastic Control, Singular Perturbation and Team Theories to an Example of Large Scale System : Management of Hydropower Production"
IEEE. A.C. April 1978.

[7] F. DELEBECQUE-J.P. QUADRAT
"Asymptotic Problems for Control of Markov Chains with Strong and Weak Interactions"
IRIA-IFAC Workshop on singular pertubations in control IRIA Publications, 1978.

[8] J.L. DERMAN
"Finite State Markovian Decision Processes"
Academic Press, New York 1970.

[9] V.G. GAITSGORI-A.A. PERVOZVANSKII
"Aggregation of States in a Markov Chain with Weak Interactions"
Kibernitika, 1975, n°5, p. 441-450.

[10] R.A. HOWARD
"Dynamic Programming and Markov Processes"
MIT, 1960.

[11] T. KATO
"Perturbation Theory for Linear Operators"
Springer Verlag, 1966.

[12] J.G. KEMENY- J.L. SNELL
"Finite Markov Chains"
Van Nostrand, Princeton 1960.

[13] M.E. LANERY
"Etude Assymptotique des Systemes Markoviens à Commandes"
RIRO, N°5, 1967, p. 3-56.

[14] B.L. MILLER-A.F. VEINOTT
"Discrete Dynamic Programming with Small Interest Rate"
Ann. Math. Stat. 1969, vol. 40, n°2, p. 366-370.

[15] R. PALLU DE LA BARRIERE
"Cours d'Automatique Théorique"
Dunod Paris, 1966.

[16] A.A. PERVOSVANSKII-I.N. SMIRNOV
"Stationary State Evaluation of a Complex System with Slow Varying coupling"
Kibernetika, 1974, n°4, p. 603- 611.

[17] U. ROTHBLUM
"Normalized Markov Decision Chains II: Optimality of Nonstationary Policies"
SIAM J. of Control and Opt. Feb. 1977, Vol. 15, N°2, p. 221-232.

[18] A.F. VEINOTT
"Discrete Dynamic Programming with Sensitive Discount Optimality Criteria"
An. Math. Statist. 40, p. 1635-1660. 1969.

[19] P.J. SCHWEITZER-A. FEDERGRUEN
"The functional equations of undiscounted Markov renewall Programming"
Math. of Op. Res. Vol. 3, N°4, November 1978.