OFFPRINT FROM

SINGULAR PERTURBATIONS IN SYSTEMS AND CONTROL

edited by

M.D. ARDEMA

CISM COURSES AND LECTURES No. 280

International Centre for Mechanical Sciences

SPRINGER-VERLAG WIEN-NEW YORK, 1983

OPTIMAL CONTROL OF PERTURBED MARKOV CHAINS: THE MULTITIME SCALE CASE.

J.P. QUADRAT

INRIA
Domaine de Voluceau - B.P. 105
78153 LE CHESNAY Cédex

ABSTRACT

Given a controled perturbed Markov chain of transition matrix $m^u(\varepsilon)$, where ε is the perturbation scale and u the control, we study the solution expansion in ε, w^ε, of the dynamic programming equation :

$$\min_u [m^u(\varepsilon)\, w^\varepsilon + c^u(\varepsilon)] = (1+\lambda(\varepsilon))w^\varepsilon .$$

$m^u(\varepsilon)$, $c^u(\varepsilon)$, $\lambda(\varepsilon)$ are polynomials in ε. The case $\lambda(\varepsilon) = \varepsilon^\ell$ leads to study Markov chains on a time scale of order $1/\varepsilon^\ell$. The state space and the control set are finite.

PLAN

1) Introduction
2) Notations and statement of the problem
3) Review of Markov chains
4) Perturbed Markov chains
5) Review of controlled Markov chains.
6) Control of perturbed Markov chains.

1. - INTRODUCTION

Stochastic or deterministic control problems can be reduced after discretization to the control of Markov chains. This approach leads to control of Markov chains which have a large number of states. An attempt to solve this difficulty is to see the initial Markov chains as the perturbation of a simpler one when this is possible. "Simple" Markov chains are Markov chains which have several recurrent classes. Then the perturbation can be seen as a small coupling between these recurrent classes. This coupling cannot be neglected on time scale of order $\frac{t}{\varepsilon}$, $\frac{t}{\varepsilon^2}$, where ε denotes the amplitude of the perturbations. Nevertheless this point of view leads to a hierarchy of more and more aggregated chains, each one being valid for a particular time scale. Their states are the recurrent classes of the faster time scale and their transition matrices can be computed explicitly. Then, in the control context, we can take advantage of this particular structure to design a faster algorithm to solve the dynamic programming equation.

This kind of problem has a long history. Gauss, for example, has studied such problems in celestian mechanics, there the recurrent classes role are played by the planet orbits. In the operations research literature studies of two time scale Markov chains has been done in Simon-Ando [12], Courtois [4], Gaitsgorï-Pervozvanski [8]. The multitime scale situation can be found in Delebecque [5], Coderch-Sastry-Willsky-Castanon [2],[3]. The two time scale control problem (actualization rate of order ε) is solved in Delebecque-Quadrat [6],[7]. The ergodic control problem when the unperturbed chain has no transient classes has been studied in Philips-Kokotovic [19]. In this paper we give the construction of the complete expansion of the optimal cost of the control problem in the general multi-time scale situation. For that we use three kinds of results :

- the Delebecque [5] result describing the reduction process of Kato [9] in the Markov chain situation.
- the realization theory of implicit systems developed by Bernhard [1]. This gives a recursive mean of computing all the cost expansion in the uncontrolled case.
- the Miller-Veinott [10] way of constructing the optimal cost expansion of an unperturbed Markov chain having a small actualization rate.

2. - NOTATIONS AND STATEMENT OF THE PROBLEM

We study the evaluation of a cost associated to the trajectory of a discrete Markov chain in four situations (unperturbed-perturbed), (controled-uncontroled); for this let us introduce some n-tuple defining completely the data of each problem, and some related notations.

2.1. - $(T, \mathcal{X}, m, c, \lambda)$ is associated to the unperturbed uncontroled case and shall be called the Markov chain n-uple.

- T is the time set isomorphic to $\mathbb{N}$;
- $\mathcal{X}$ is the state space of the Markov chain, is a finite discrete space. $|\mathcal{X}|$ denotes card($\mathcal{X}$) that is the number of elements of $\mathcal{X}$. x will be the generic element of $\mathcal{X}$;
- m is the transition matrix of the Markov chain, that is a $(|\mathcal{X}|, |\mathcal{X}|)$-matrix with positive entries such that $\sum_{x' \in \mathcal{X}} m_{x x'} = 1$;
- c is the instantaneous cost that is a $|\mathcal{X}|$-vector with positive entries;
- λ is an actualization rate that is, $\lambda \in R$ and $\lambda > 0$.

The set of possible trajectories is denoted by $\Omega = \mathcal{X}^T$, a trajectory by $\omega \in \Omega$, the position of the process at time t if the trajectory is ω by $X(t,\omega)$. The conditional probability of the cylinder :

$$B = \{\omega : X_t(\omega) = x_t,\ t = 0,1,\ldots,n\}$$

knowing $X(0,\omega)$ is :

$$P^{X_0}(B) = \prod_{t=0}^{n-1} m_{x_t x_{t+1}}$$

To the trajectory ω is associated the cost :

$$j(\omega) = \sum_{t=0}^{+\infty} \frac{1}{(1+\lambda)^{t+1}} c_{X(t,\omega)} \tag{2.1}$$

The conditional expected cost knowing $X(0,\omega)$ is a $|\mathfrak{X}|$-vector denoted w defined by :

$$w_x := E[j(\omega) \mid X(0,\omega) = x], \forall x \in \mathfrak{X} \tag{2.2}$$

The Hamiltonian is the operator :

$$h : \mathbb{R}^{|\mathfrak{X}|} \to \mathbb{R}^{|\mathfrak{X}|} \quad w \quad [m-(1+\lambda)i]w+c \tag{2.3}$$

where i denotes the identity of the set of $(|\mathfrak{X}|,|\mathfrak{X}|)$-matrices.

Then w defined by (2.2) is the unique solution of the Kolmogorov equation :

$$h(w) = 0 \tag{2.4}$$

2.2. – In the perturbed situation the n-tuple defining the perturbed Markov chain is :

$$(T, \mathfrak{X}, \mathcal{E}, m(\varepsilon), c(\varepsilon), \lambda(\varepsilon))$$

– $\mathcal{E}$ is now the space of the perturbations ; in all the following it $\mathbb{R}^+$;

– $m(\varepsilon)$, $c(\varepsilon)$, $\lambda(\varepsilon)$ have the same definition as previously but depends on the parameter $\varepsilon \in \mathcal{E}$, and we suppose that they are polynomials in this variable.

We denote by d° the degree of a polynomial and by v its valuation (the smallest non zero power of the polynominal). In the following $d^o(m) = 1$, $v(m)=0$, $v(\lambda) = v(c) = \ell$.

The Hamiltonian of the perturbed problem is denoted by :

$$h(w,\varepsilon) = [m(\varepsilon) - (1+\lambda(\varepsilon))i]W + c(\varepsilon) \tag{2.5}$$

The expected conditional cost is denoted w^ε and is solution of the Kolmogorov equation :

$$h(w^\varepsilon,\varepsilon) = 0 \tag{2.6}$$

We shall prove that w^ε admits an expansion in ε that we shall denote by $W(\varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n w_n$ where W_i are $|\mathcal{X}|$-vectors; then we have :

$$m(\varepsilon)\ W(\varepsilon) = \sum_{n=0}^{\infty} \varepsilon^n (MW)_n \tag{2.7}$$

with :

$$M = \begin{bmatrix} m_0 & & & 0 \\ m_1 & m_0 & & \\ 0 & m_1 & m_0 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \tag{2.8}$$

an infinite block matrix.

For the Hamiltonian we can introduce the same notation :

$$h(W(\varepsilon),\varepsilon) = \sum_{n}^{\infty} \varepsilon^n H_n(W) \tag{2.9}$$

where $H_n(W)$ are the $|\mathcal{X}|$-vectors defined in (2.9) by identification of the ε^i terms; that is :

$$\begin{cases} H_0(W) = (m_0-i)w_0 \\ H_1(W) = m_1\, w_0 + (m_0-i)w_1 \\ \vdots \\ H_\ell(W) = -\lambda_\ell\, w_o + m_1\, w_{\ell-1} + (m_0-i)w_\ell + c_\ell \\ \vdots \end{cases} \tag{2.10}$$

(2.10) can written H(W) with :

$$H(W) \equiv [M - (I+\Lambda)]W + C, \tag{2.11}$$

where :

$C = (c_n,\ n \in \mathbb{N},\ c_n$ are $|\mathcal{X}|$-vectors)

I : the identity operator $\begin{bmatrix} i & 0 & 0 & 0 & \dots \\ 0 & i & 0 & 0 & \dots \\ 0 & 0 & i & & \\ \vdots & \vdots & & \ddots & \end{bmatrix}$

Λ : the operator ℓ^{th} $|\mathcal{X}|$-block $\begin{bmatrix} & & 0 \\ i\lambda_\ell & \ddots & \\ 0 & \ddots & \end{bmatrix}$

An expansion of the cost is obtained by solving :

$$H(W) = 0 \tag{2.12}$$

Moreover the sequence $(W_i,\ i \in \mathbb{N})$ can be computed recursively. These two results will be shown in part 4.

2.3. - For the control problem we need the introduction of the n-tuple:

$(T, \mathcal{X}, \mathcal{U}, m^u, c^u, \lambda)$

- $\mathcal{U}$ is the set of control which is here a finite set ; $|\mathcal{U}|$ denotes the cardinal of $\mathcal{U}$; its generic element is denoted by u ;
- m denotes the $(|\mathcal{U}|, |\mathcal{X}|, |\mathcal{X}|)$ tensor of entries $m^u_{xx'}$, the probability to

go in x', starting from x, the control being u.

- c denotes the $(|\mathcal{U}|, |\mathcal{X}|)$ matrix of entries c_x^u, the cost to be in x, the control being u.

A policy is an application :

$$s : \mathcal{X} \to \mathcal{U}.$$

The set of policies is $\mathcal{P} := \mathcal{U}^{\mathcal{X}}$.

For a policy s, mos denotes the $(|\mathcal{X}|, |\mathcal{X}|)$ transition matrix of entries :

$$(mos)_{xx'} = m_{xx'}^{s_x} ; \tag{2.13}$$

cos denotes the $|\mathcal{X}|$-vector :

$$(cos)_x = c_x^{s_x}. \tag{2.14}$$

We associate to a policy $s \in \mathcal{P}$ and a trajectory ω, the cost

$$j^s(\omega) = \sum_{t=0}^{+\infty} \frac{1}{(1+\lambda)^{t+1}} (cos)_{X(t,\omega)} \tag{2.15}$$

and the optimal conditional expected cost knowing the initial condition is :

$$w_x^* = \underset{s \in \mathcal{P}}{\mathrm{Min}} \; \mathbb{E}(j^s(\omega) \mid X(0,\omega) = x) \tag{2.16}$$

The Hamiltonian is defined as the operator :

$$h : \mathcal{U} \times \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X}} \tag{2.17}$$
$$(u,w) \qquad h^u(w) = [m^u - (1+\lambda)i]w + c^u.$$

The notation $(hos)_x$ for $h_x^{s_x}$ will be used.

Then the optimal Hamiltonian is the operator :

$$h^* : \mathbb{R}^{\mathcal{X}} \to \mathbb{R}^{\mathcal{X}} \qquad w \quad h^*_x(w) = \min_u h^u_x(w), \forall x \in \mathcal{X} \tag{2.18}$$

The optimal expected cost w^* is the unique solution of the dynamic programming equation :

$$h^*(w^*) = 0 \tag{2.19}$$

An optimal policy is given by :

$$s^* : \mathcal{X} \to \mathcal{U} \qquad x \quad s^*_x \in \operatorname{argmin} h^u_x(w^*), \forall x \in \mathcal{X}.$$

2.4. - The perturbed control problem is defined by the n-tuple:

$$(T, \mathcal{X}, \mathcal{U}, \mathcal{E}, m^u(\varepsilon), c^u(\varepsilon), \lambda(\varepsilon)).$$

Its interpretation is clear from the previous paragraphs.

By analogy the notations $H^u(w,\varepsilon)$, $h^*(w,\varepsilon)$, $w^{*\varepsilon}$, $H^u(W)$ are clear, but we need a definition of $H^*(W)$. For that let us introduce the lexicographic order, $\succcurlyeq$, for sequences of real numbers, that is :

$$(y_0, y_1, \ldots) \succcurlyeq (y'_0, y'_1, y'_2, \ldots) \text{ is true} \iff (\text{if } y_n = y'_n, \forall n < m \text{ then } y_m \geq y'_m) \ \forall m \in \mathbb{N}. \tag{2.20}$$

We denote by $\vec{\min}$ the minimum for this order. Then we define H^* by :

$$H^*_{.x}(W) = \vec{\min}\, H^u_{.x}(W) \tag{2.21}$$

(indeed $H^u_{.x}(W)$ is a sequence of real numbers).

We shall prove that $w^{*\varepsilon}$ admits an expansion in ε denoted by $W^*(\varepsilon)$ which satisfies :

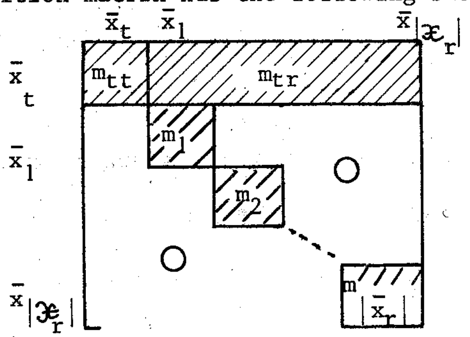
$$H^*(W^*) = 0 \tag{2.22}$$

The purpose of this paper is to prove this last result and to show that W^* can be computed recursively. By this way we can design faster algorithm than the ones obtained by a direct solution of $h^*(w^{*\varepsilon},\varepsilon) = 0$.

3. - REVIEW OF MARKOV CHAINS

Let us recall some facts on Markov chains. We consider the Markov chain defined by $(T,\mathcal{X}, m, c, \lambda)$.

The matrix m defines a connexity in the state space $\mathcal{X}$, that is : $x \in \mathcal{X}$ and $x' \in \mathcal{X}$ are connected if there exists a non zero probability path between x and x'. Moreover if x' and x are also connected we say that x and x' are strongly-connected. The equivalence classes of the strongly-connexity relation defines a partition on $\mathcal{X}$. The connexity relation defines a partial order on these classes. The final classes for this order are the recurrent classes of the Markov chain. Their set is denoted by $\bar{\mathcal{X}}_r = \{\bar{x}_1,\bar{x}_2,\dots,\bar{x}_{|\bar{\mathcal{X}}_r|}\}$. The other states are called transient and their set is denoted by $\bar{x}_t$. Thus we have defined a partition $\bar{\mathcal{X}}$ of $\mathcal{X}$, $\bar{\mathcal{X}} = \bar{x}_t \cup \bar{\mathcal{X}}_r$. Let us consider the natural numerotation of the states of $\mathcal{X}$ after the grouping defined by the partition $\bar{\mathcal{X}}$. With this numeration the transition matrix has the following block structure :



m_{tr} has at least a non zero term by row (3.1)

m admits the eigenvalue 1 because $\sum_{x'\in\mathcal{X}} m_{xx'} = 1$. This eigenvalue is semi-simple (the eigen-space associated to the eigenvalue 1 admits a base of eigenvectors). This can be proved easily by remarking that $|m|_{\infty,\infty} = 1$ which proves that $|m^n|_{\infty,\infty} = 1$, where $|m|_{\infty,\infty}$ denotes the norm of matrices seen as operators on $\mathbb{R}^{|\mathcal{X}|}$ with the sup norm. Now if the eigenvalue 1 was not semi-simple, in an appropriate basis m would have a Jordan block : $\begin{bmatrix} 1 & 1 & 0 \\ & 1 & 1 \\ 0 & & 1 \end{bmatrix}$ and we should have $|m^n|_{\infty,\infty} \xrightarrow[n\to\infty]{} \infty$ which is a contradiction. From this property we see that we have the decomposition :

$$\mathbb{R}^{\mathcal{X}} = \mathcal{N}(a) \oplus \mathcal{R}(a) \tag{3.2}$$

where :

$$a \equiv m - i \tag{3.3}$$

$\mathcal{N}(a)$ denotes the kernel of $\mathcal{R}(a)$, the range of the operator a.

To define the projector a^0 on $\mathcal{N}(a)$ parallel to $\mathcal{R}(a)$ we need to know a base of $\mathcal{N}(a)$ and $\mathcal{N}(a')$ where ' denotes the transposition.

The set $\{p_{\bar{x}.}, \bar{x} \in \bar{\mathcal{X}}_r\}$ of the extremal invariant probability measures of m defines a base of $\mathcal{N}(a')$. $p_{\bar{x}}$ has for support $\bar{x}$ and the restriction to $\bar{x}$ of $p_{\bar{x}.}$ denoted by $\bar{p}_{\bar{x}}$ satisfies :

$$\bar{p}_{\bar{x}}\, m_{\bar{x}} = \bar{p}_{\bar{x}} \tag{3.4}$$

This result is clear from (3.1).

The set $\{q_{.\bar{x}}, \bar{x} \in \bar{\mathcal{X}}_r\}$ where $q_{x\bar{x}}$ denotes the probability starting from x to end in $\bar{x}$ defines a base of $\mathcal{N}(a)$. Indeed $q_{\bar{x}}$ satisfies :

$$q_{x\bar{x}} = \begin{cases} 0 & \text{if } x \notin \bar{x} \cup \bar{x}_t \\ 1 & \text{if } x \in \bar{x} \\ \bar{q}_{\bar{x}} & \text{if } x \in \bar{x}_t \end{cases} \tag{3.5}$$

with $\bar{q}_{\bar{x}}$ solution of the Dirichlet problem :

$$m_{tt}\,\bar{q}_{\bar{x}} = -\,m_{t\bar{x}}\,1^{*} \tag{3.6}$$

From (3.5) and (3.6) it is clear that $q_{\bar{x}} \in \mathcal{N}(a)$, $(q_{\bar{x}}, \bar{x} \in \bar{\mathcal{X}}_r)$ are linearly independent and from (3.1) that they form a base of $\mathcal{N}(a)$.

If we see p as $(|\bar{\mathcal{X}}|, |\mathcal{X}|)$-matrix and q as a $(|\mathcal{X}|, |\bar{\mathcal{X}}|)$-matrix the projector on $\mathcal{N}(a)//\mathcal{R}(a)$ is :

$$a^{0} = q\,p. \tag{3.7}$$

We have :

$$aa^{0} = a^{0}a = 0 \tag{3.8}$$

There exists a pseudo inverse a^{+} of $-a$ which is the inverse of $-a$, restricted to $\mathcal{R}$ (a), defined precisely by the relations:

$$\begin{cases} a^{+}a = a^{+}a = a^{0} - i \\ a^{+}a^{0} = a^{0}a^{+} = 0 \end{cases} \tag{3.9}$$

τ is a random time, that is a random variable on T, independent of the Markov chain X_t, of exponentiel probability law of parameter λ that is :

$$P(\tau=t) = \frac{\lambda}{(1+\lambda)^{t+1}} \tag{3.10}$$

We have :

$$\mathbb{E}(\tau) = \frac{1}{\lambda} \tag{3.11}$$

On the new probability space $\tilde{\Omega} = \Omega \otimes T$ we have :

---

* 1 denotes the $|\bar{x}|$-vector : $1_x = 1$, $\forall x \in \bar{x}$.

$$\lambda\, w_x = \mathbb{E}\,[\sum_{t=0}^{+\infty} \frac{\lambda}{(1+\lambda)^{t+1}}\, c_{X(t,\omega)} \mid X(0,\omega) = x]$$
$$= \mathbb{E}[c_{X(\tau(\tilde\omega),\tilde\omega)} \mid X(0,\tilde\omega) = x] \tag{3.12}$$

The operator :

$$r_\lambda(a) : \mathbb{R}_c^{\mathcal{X}} \to \mathbb{R}_w^{\mathcal{X}} \tag{3.13}$$

is called the resolvent of a.

From (3.12) we see that λr_λ defines a transition matrix :

$$[\lambda r_\lambda(a)]_{xx'} = \mathbb{P}\,\{X[\tau(\tilde\omega),\tilde\omega] = x' \mid X(0,\tilde\omega) = x\},\ \forall x,x' \in \mathcal{X} \tag{3.14}$$

which corresponds to the initial Markov seen at random time $\tau_1, \tau_2, \ldots, \tau_n$ with $\tau_{i+1} - \tau_i$ independent of τ_i and having the same probability law as τ .

From the Jordan form of a and the previous discussion on the semi-simple nature of the eigenvalue 1 of m, we can show the ergodic theorem :

$$\lim_{\lambda\to 0} \lambda r_\lambda(a) = a^0. \tag{3.15}$$

4. - PERTURBED MARKOV CHAIN

We study the perturbed Markov chain $(T, \mathcal{X}, \mathcal{E}, m(\varepsilon), c(\varepsilon), \lambda(\varepsilon))$, in the case $\lambda = \varepsilon^\ell \mu$, $v(c) = \ell$; that is, we study the transfer function $\varepsilon^\ell \mu(\varepsilon^\ell \mu + i - m(\varepsilon))^{-1}$ in ε. With the interpretation (3.14), this means that we look at the Markov chain on the time scale $\frac{t}{\varepsilon^\ell}$; for time scale interpretation in time domain see also Coderch-Sastry-Willsky-Castanon [2],[3]. We have seen in (2.11) that when the optimal conditional expected cost w^ε admits an expansion, $W(\varepsilon)$, in ε this expansion satisfies :

$$H(W) \equiv (M - I - \Lambda)W + C = 0 \tag{4.1}$$

(4.1) is an infinite set of linear equations. Conversely if a solution of

(4.1) exists with for example $(W_i,\ i \in \mathbb{N})$ bounded then $W(\varepsilon)$ converges, for $\varepsilon < 1$, and is a solution of :

$$h(w,\varepsilon) = 0 \tag{4.2}$$

Let us show now that (4.1) can be computed recursively.

For that we build the implicit realization of W :

$$\begin{cases} E\, y_{n+1} = F\, y_n - G\, C_{n+\ell+1} \\ W_{n+1} = H\, y_{n+1} \end{cases} \qquad y_{-1} = 0 \tag{4.3}$$

with :

$$a_0 = m_0 - i \tag{4.4}$$

$$E = \left[\begin{matrix} a_0 & & & 0 \\ m_1 & \ddots & & \\ & \ddots & \ddots & \\ -\mu\ \ 0 & & m_1 & a_0 \end{matrix}\right] \quad (\ell+1) \text{ blocks} \tag{4.5}$$

$$F = \left[\begin{matrix} 0 & a_0 & & 0 \\ \vdots & m_1 & \ddots & \\ \vdots & 0 & m_1 & a_0 \\ 0 & \cdots & \cdots & 0 \end{matrix}\right] \quad (\ell+1) \text{ blocks} \tag{4.6}$$

$$G = \left[\begin{matrix} 0 \\ \vdots \\ 0 \\ i \end{matrix}\right] \quad (\ell+1) \text{ blocks} \tag{4.7}$$

$$H = \underbrace{[\,i \quad 0 \ \cdots \ 0\,]}_{(\ell+1)\text{ blocks}} \tag{4.8}$$

Indeed if W is a solution of (4.1) :

$$y_n = (W_n, W_{n+1}, \dots, W_{n+\ell})$$

is a solution of (4.3).

Conversely if W is a solution of (4.3), by elimination of the variables y we see that W satisfies (4.1).

To prove the existence of a solution of (4.3), following Bernhard [1] we have to show that there exists $\mathcal{Z} \subset \mathbb{R}^{|\mathcal{X}| \times (\ell+1)}$ which satisfies :

$$F\mathcal{Z} \subset E\mathcal{Z}, \tag{4.9}$$

$$G \subset E\mathcal{Z}. \tag{4.10}$$

We can take $\mathcal{Z} = \mathbb{R}^{|\mathcal{X}| \times (\ell+1)}$. Indeed (4.9) is equivalent to finding a $z \in \mathbb{R}^{\mathcal{X} \times (\ell+1)}$ such that:

$$Ez = Fy \;;\; \forall y \in \mathbb{R}^{|\mathcal{X}| \times (\ell+1)}. \tag{4.11}$$

But by the change of variables $z'^k = z^k - y^{k+1}$, (4.11) becomes :

$$Ez' = Gc \text{ with } c = -\mu\, y^2 + a_1\, y^\ell \in \mathbb{R}^{\mathcal{X}} \tag{4.12}$$

which is a relation of (4.10) kind.

Delebecque [5] has proved that (4.10) has a solution. Let us show this result in two cases $\ell = 1$ and $\ell = 2$; then the general proof can be induced easily.

$\ell = 1$

We have to solve :

$$\begin{cases} a_0\, W_0 = 0 \\ (m_1-\mu)W_0 + a_0\, W_0 = -C_1. \end{cases} \tag{4.13}$$

(4.13) implies :

$$a_0^0 W_0 = W_0 \tag{4.14}$$

where a_0^0 denotes the projector on $\mathcal{N}(a_0) // \mathcal{R}(a_0)$.

Then left multiplying (4.13) by a_0^0 gives :

$$a_0^0(\mu - m_1) a_0^0 W_0 = a_0^0 C_1 \tag{4.15}$$

Using the factorization of $a_0^0 = qp$ where q and p are defined as in (3.4) and (3.5) for $m = m_0$ we have :

$$W_0 = q(\mu - p\, m_1\, q)^{-1}\, p\, C_1 \qquad \forall \mu \notin \text{Spect.}(p\, m_1\, q) \tag{4.16}$$

Thus :

$$W_1 = a_0^+ [(m_1 - \mu) W_0 + C_1] + \bar{W}_1 \quad \forall \bar{W}_1 : \bar{W}_1 = a_0^0 \bar{W}_1 \tag{4.17}$$

where a_0^+ denotes the pseudo inverse of $-a_0$ defined in (3.9).

Then we have proved that W_0 is defined uniquely and W_1 up to an element of $\mathcal{N}(a_0)$.

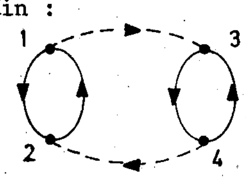
From the stochastic interpretation :

$$W_0 = \lim_{\varepsilon \to 0} \mu\, \varepsilon (1 + \mu\, \varepsilon - m(\varepsilon))^{-1} \frac{C(\varepsilon)}{\mu\, \varepsilon}, \tag{4.18}$$

it follows that $p\, m_1\, q$ is a generator of a Markov chain and thus that $(\mu - p\, m_1\, q)^{-1}$ exists $\forall \mu > 0$.

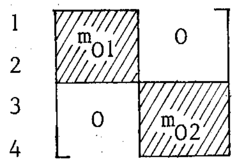
Example

Consider the Markov chain :



where the dotted lines corresponds to probabilities of order ε, the other lines to order-1 ones.

m_0 has the following block structure :



The dimension of $\mathcal{N}(a_0)$ is 2.

$$q = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \tag{4.19}$$

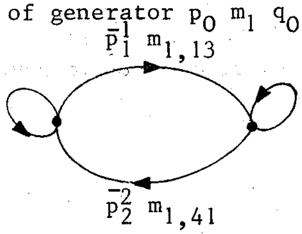
$$p = \begin{bmatrix} \bar{p}_1^1 & \bar{p}_1^2 & 0 & 0 \\ 0 & 0 & \bar{p}_2^1 & \bar{p}_2^2 \end{bmatrix} \tag{4.20}$$

with :

$$\bar{p}_1 \; m_{01} = \bar{p}_1 \tag{4.21}$$

$$\bar{p}_2 \; m_{02} = \bar{p}_2 \tag{4.22}$$

The aggregated chain of generator $p_0 \; m_1 \; q_0$ is :



$\underline{\ell = 2}$

We have to solve :

$$\begin{cases} a_0 \; W_0 = 0 \\ m_1 \; W_0 + a_0 \; W_1 = 0 \\ -\mu \; W_0 + m_1 \; W_1 + a_0 \; W_2 = - C_2 \end{cases} \tag{4.23}$$

Left multiplying (4.23) by a_0^0 we have :

$$\left\{\begin{array}{l} a_0^0\, W_0 = W_0 \\ a_0^0\, m_1\, a_0^0\, W_0 = 0 \qquad W_1 = a_0^+\, m_1\, W_0 + a_0^0\, W_1 \\ a_0^0\, m_1\, W_1 - \mu\, a_0^0\, W_0 = -\, a_0^0\, C_2 \end{array}\right. \tag{4.24}$$

which gives using the formula of W_1 :

$$\left\{\begin{array}{l} a_0^0\, W_0 = W_0 \\ a_0^0\, m_1\, a_0^0\, W_0 = 0 \\ a_0^0\, m_1\, a_0^0\, W_1 + a_0^0\, m_1\, a_0^+\, m_1\, a_0^0\, W_0 - \mu\, a_0^0\, W_0 = -a_0^0\, C_2 \end{array}\right. \tag{4.25}$$

Using the notations :

$$\left\{\begin{array}{l} \bar{a}_0 = p\, m_1\, q \\ \bar{m}_1 = p\, m_1\, a_0^+\, m_1\, q \\ \bar{W}_0 = p\, W_0 \\ \bar{W}_1 = p\, W_1 \\ \bar{C}_1 = a_0^0\, C_2 \end{array}\right. \tag{4.26}$$

(4.25) becomes:

$$\left\{\begin{array}{l} \bar{a}_0\, \bar{W}_0 = 0 \\ (\bar{m}_1 - \mu)\bar{W}_0 + \bar{a}_0\, \bar{W}_1 = -\bar{C}_1 \end{array}\right. \tag{4.27}$$

(4.26) is a problem of kind $\ell = 1$. Thus using the factorization of $\bar{a}_0^0 = \bar{p}\,\bar{q}$ which exists because $\bar{a}_0$ is a generator of a Markov chain (see $\ell = 1$ case), we obtain :

$$W_0 = q\,\bar{q}\,(\mu - \bar{p}\,\bar{m}_1\,\bar{q})^{-1}\,\bar{q}\,q\,C_2 \quad \forall \mu \notin \text{Spect.}(\bar{p}\,\bar{m}_1\,\bar{q}) \tag{4.28}$$

Then from (4.27) and (4.24) we can compute W_1 and W_2. W_1 is defined up to an element of $q\mathcal{N}(\bar{a}_0)$, W_2 up to an element of $\mathcal{N}(a_0)$.

The stochastic interpretation of W_0 :

$$W_0 = \lim_{\varepsilon\to 0} \mu\varepsilon^2 \,(1+\varepsilon^2\mu - m(\varepsilon))^{-1} \frac{C(\varepsilon)}{\mu\,\varepsilon^2} \tag{4.29}$$

shows that $\bar{p}\,\bar{m}_1\,\bar{q}$ is a generator of a Markov chain. Thus we have the existence of (4.21), $\forall\mu > 0$. ∎

This procedure can be reiterated and gives the general case (Delebecque [5]). In this reference we find the relation of this method and the reduction process of Kato [9]. ∎

The reiteration of the reduction process finishes when the aggregate chain obtained has the same number of recurrent classes as the one of the initial chain $(m(\varepsilon))$. ∎

Bernhard [1] has proved that the solution of the implicit system is unique if $\mathcal{N}(E) \cap \mathcal{Z} = \emptyset$. The discussion of the two cases $\ell = 1$ and $\ell = 2$ shows that $\mathcal{N}(E) \cap \mathbb{R}^{|\mathcal{X}|\times(\ell+1)} \neq 0$ in the case $\ell = 1$:

$$\mathcal{N}(E) \cap \mathbb{R}^{2|\mathcal{X}|} = \mathcal{N}\left(\begin{bmatrix} i & 0 \\ 0 & a_0 \end{bmatrix}\right), \tag{4.30}$$

In the case $\ell = 2$, on an appropriate basis :

$$\mathcal{N}(E) \cap \mathbb{R}^{3|\mathcal{X}|} = \mathcal{N}\left(\begin{bmatrix} i & & & 0 \\ & i_0 & & \\ & 0 & \bar{a}_0 & \\ & & & a_0 \end{bmatrix}\right) \tag{4.31}$$

where i_0 denotes the identity on $\mathcal{R}(a_0)$.

But we see that the non unicity part of the implicit system is inobservable in the output; indeed $\mathcal{N}(H) \supset \mathcal{N}(E)$. This property is true in the general situation because we can prove that W_0 is always defined uniquely.

We have proved the :

Theorem 1 : The solution W^{ε} of :

$$h(W,\varepsilon) := (m(\varepsilon) - i - \lambda(\varepsilon))W + c(\varepsilon) = 0, \tag{4.32}$$

admits an expansion $W(\varepsilon)$ which is the unique solution of :

$$H(W) := (M-I-\Lambda)W + C = 0 \tag{4.33}$$

Moreover W can be computed recursively by solving the implicit system realization of (4.36) :

$$\begin{cases} Ey_{n+1} = Fy_n - GC_{n+\ell+1}, \; y_{-1} = 0, \\ W_{n+1} = Hy_{n+1}, \end{cases} \tag{4.34}$$

where E,F,G,H are defined in (4.5) to (4.8).

This implicit system has an output uniquely defined and it admits a strictly causal realization.

The first term of the expansion has the interpretation of the conditional expected cost of an aggregated Markov chain obtained by reiteration of an aggregation procedure which consists in aggregating the recurrent classes of the order 1 transition matrix, in one state, and computing the transition matrix of the new aggregate chain.

5. - REVIEW OF CONTROLLED MARKOV CHAINS

Given the controlled Markov chain n-tuple: $(T, \mathcal{X}, \mathcal{U}, m^u, c^u, \lambda)$. The optimal conditional expected w^* cost is the unique solution in w of the dynamic programming equation :

$$h_x^*(w) \equiv \min_u [(m^u - 1 - \lambda)w + c^u]_x = 0, \; \forall x \in \mathcal{X}. \tag{5.1}$$

This result can be proved using the Howard algorithm :

Step 1 : Given a policy $s \in \mathcal{U}^{\mathfrak{X}}$, let us compute w, solving, in w, the linear equation :

$$h o s(w) = 0 \tag{5.2}$$

Step 2 : Given a conditional expected cost w, let us improve the policy by computing :

$$\min_u h_x^u(w) \tag{5.3}$$

We change s(x) only if $h_x^u(w) < 0$. Then we return to step 1.

By this way we generate a sequence :

$$((s^n, w^n) \ ; \ n \in \mathbb{N})$$

which converges after a finite number of steps. The sequence $(w^n, n \in \mathbb{N})$ is decreasing.

Indeed :

$$h o s^n(w^n) = 0 \tag{5.4}$$

$$h o s^{n+1}(w^{n+1}) = 0 \tag{5.5}$$

Then (4.4)-(4.5) gives :

$$(m o s^{n+1} - 1 - \lambda)(w^n - w^{n+1}) + h o s^n(w^n) - h o s^{n+1}(w^n) = 0 \tag{5.6}$$

But by (4.3) we have :

$$h o s^n(w^n) - h o s^{n+1}(w^n) \geq 0 \tag{5.7}$$

Then (5.6) and (5.7) proves that :

$$w_n - w_{n+1} \geq 0 \tag{5.8}$$

Indeed, (5.6) can be seen as a Kolmogorov equation in $(w_n - w_{n+1})$, with a positive instantaneous cost.

The existence and the uniqueness of a solution in w of (4.1) follows easily from this result.

6. - CONTROL OF PERTURBED MARKOV CHAINS

Given the perturbed controlled Markov chain n-tuple $(T, \mathcal{X}, \mathcal{U}, \mathcal{E}, m^u(\varepsilon), c^u(\varepsilon), \lambda(\varepsilon))$. The optimal cost is the unique solution in w of the dynamic programming equation :

$$h^*_x(w,\varepsilon) \equiv \min_u[(m^u(\varepsilon)-1-\lambda(\varepsilon))w + c^u(\varepsilon)]_x = 0, \forall x \in \mathcal{X} \tag{6.1}$$

We have the :

Theorem 2 : The solution of (6.1) denoted by $w^{*\varepsilon}$ admits an expansion in ε denoted by $W^*(\varepsilon)$ which is the unique solution in W of the vectorial dynamic programming equation :

$$H^*_{.x}(W) \equiv \overrightarrow{\min_u}[(M^u-I-\Lambda)W + C^u]_{.x} = 0, \forall x \in \mathcal{X} \tag{6.2}$$

Let us remember that $\overrightarrow{\min}$ means the minimum for the lexicographic order on the sequence of real numbers.

The solution W^* can be computed by the vectoriel Howard algorithm :

Step 1 : Given a policy $s \in \mathcal{U}^{\mathcal{X}}$, let us compute W using the results of part 4 :

$$H^s_o(w) = 0 \tag{6.3}$$

Step 2 : Given a conditional expected cost W, let us improve the policy by computing :

$$\overrightarrow{\mathrm{Min}_u}\, H^u_{.x}(W) \tag{6.4}$$

We change s(x) only if $H^u_{.x}(W) \prec 0$. Then we return to step 1.

By this way we generate a sequence :

$$((s^n, W^n) \; ; \; n \in \mathbb{N})$$

which converges after a finite number of steps. The sequence $(W^n, \; n \in \mathbb{N})$ is decreasing for the lexicographic order $\succ$.

This decreasing property can be proved easily using the corresponding proof in the unperturbed case parts, and the following equivalence :

$$h_x^u(W(\varepsilon),\varepsilon) \geq h_x^{u'}(W(\varepsilon),\varepsilon) \iff H^u_{.x}(W) \not\succ H^{u'}_{.x}(W). \qquad (6.5)$$

From this property the theorem can be proved easily.

A priori it is not clear if we may restrict the minimization to finite part of the infinite sequence.

The following result shows that this is possible and gives an estimate on the length of the sequence part on which we have to apply the lexicographic order minimization.

<u>Theorem 3</u> : <u>The vectoriel minimization (in 6.4) may be applied on the $\eta = (d^o(C) + (v(\lambda) + 2)|\mathcal{X}|)$ first terms of the sequence only without changing the convergence to the solution of th. 2.</u>

<u>Proof</u> : Let us show that :

$$H_n^u(W) = H_n^{u'}(W), \; \forall i = d^o(c)+1,\ldots,\eta \implies H_n^u(W) = H_n^{u'}(W), \forall \eta > d^o(c) \qquad (6.6)$$

By theorem 1, W admits a strictly causal realization that is there exists $\tilde{E}$, $\tilde{F}$, $\tilde{G}$ such that :

$$\begin{cases} z_{n+1} = \tilde{E}\, z_n + \tilde{F}\, C_{n+\ell+1} \\ W_{n+1} = \tilde{G}\, z_{n+1} \end{cases} \qquad (6.7)$$

For that we have equal to zero the control corresponding to the non-unicity of the implicit system (4.34) because this non-unicity is not observable. By (4.34) we know that the order of the matrix E is smaller than $(v(\lambda)+1)|\mathfrak{X}|$. The entry $C_{n+\ell+1}$ is equal to zero for $n \geq d^0(c)-v(c)$.

We add $|\mathfrak{X}|$ new states to z, denoted by $\tilde{z}$ with :

$$\tilde{z}_{n+1} = W_n. \tag{6.8}$$

With the new state $\check{z} = (z,\tilde{z})$ the second part of (6.6) can be written :

$$([a_0^u - a_0^{u'}][\tilde{G},0] + [m_1^u - m_1^{u'}][0,i])\ \check{z}_n = 0 \tag{6.9}$$

(6.9) has the form :

$$J\ \check{z}_n = 0 \tag{6.10}$$

with J an observation matrix of the dynamical system of state $\check{z}_n$. It follows by the Cayley-Hamilton theorem that if (6.10) is true $\forall n : \eta \geq n > d^0(c)$ then (6.10) is true $\forall n > d^0(c)$. The theorem 3 is deduced easily from this result.

Remark :

The order of E in 4.34 is $(v(\lambda)+1)|\mathfrak{X}|$ but the order of $\tilde{E}$ is much smaller. It is certainly of order $|\mathfrak{X}|$. Moreover it is certainly not necessary to memorize completely V_n in (6.8) to be able to compute it from z_{n+1}, thus the value of η is much smaller than the value given in the theorem 3.

Stochastic interpretation of the first term of the expansion can be found in Delebecque - Quadrat [6], [7].

REFERENCES

[1] P. BERNHARD. Sur les systèmes dynamiques linéaires implicites singuliers, à paraître SIAM J. on control and optimization et rapport INRIA 69, 1981.

[2] M. CODERCH, A.S. WILLSKY, S.S. SASTRY, D.A. CASTANON. Hierarchical aggregation of linear systems with multiple time scales, MIT Report LIDS-P-1187, mars 1982.

[3] M. CODERCH, A.S. WILLSKY, S.S. SASTRY. Hierarchical aggregation of singularly perturbed finite state Markov chains submitted to stochastics.

[4] P.J. COURTOIS. Decomposability, ACM Monograph Series, Academic Press, 1977.

[5] F. DELEBECQUE. A reduction process for perturbed Markov chains, à paraître SIAM J. of applied math. to appear.

[6] F. DELEBECQUE, J.P. QUADRAT. Optimal control of Markov chains admitting strong and weak interactions, Automatica, Vol. 17, n° 2, pp. 281-296, 1981.

[7] F. DELEBECQUE, J.P. QUADRAT. The optimal cost expansion of finite controls finite states Markov chains with weak and stong interactions. Analysis and optimization of systems, Lecture Notes an control and Inf. Science 28 Springer Verlag, 1980.

[8] A.A. PERVOZVANSKII, A.V. GAITSGORI. Decomposition aggregation and approximate optimization en Russe, Nauka, Moscou, 1979.

[9] T. KATO. Perturbation theory for linear operator, Springer Verlag, 1976.

[10] B.L. MILLER, A.F. VEINOTT. Discrete dynamic programming with small interest rate. An. math. stat. 40, 1969, pp. 366-370.

[11] R. PHILIPS, P.KOKOTOVIC . A singular perturbation approach to modelling and control of Markov chains IEEE A.C. Bellman issue, 1981.

[12] H. SIMON, A. ANDO. Aggregation of variables in dynamic systems, Econometrica, 29, 111-139, 1961 .

[13] J. KEMENY, L. SNELL. Finite Markov chains, Van Nostrand, 1960.

[14] O. MURON. Evaluation de politiques de maintenance pour un système complexe, RIRO, vol. 14, n° 3, pp. 265-282, 1980.

[15] S.L. CAMPBELL, C.D. MEYER Jr. Generalized inverses of linear transformations. Pitman, London, 1979.