NUMERICAL SOLUTION OF THE OPERATOR RICCATI EQUATION FOR THE FILTERING OF LINEAR STOCHASTIC HEREDITARY DIFFERENTIAL SYSTEMS*

M.C. Delfour

Centre de Recherches Mathématiques
Université de Montréal, C.P. 6128,
Montréal, P.Q. H3C 3J7 Canada.

1. Introduction.

The filtering problem for hereditary systems has been considered by a number of authors. To the author's knowledge, the first paper in this field is the one of H. KWAKERNAAK [1], where simultaneously the smoothing and filtering problems for linear differential systems with multiple constant time delays are studied. Other more recent papers by A. LINDQUIST [1], A. BENSOUSSAN [2], BENSOUSSAN-DELFOUR-MITTER [1], MITTER-VINTER [1], R. CURTAIN [1] and R. KWONG [1] have also discussed the theory of this problem and extended the well-known duality theorem of KALMAN-BUCY in various forms.

In this paper we put our hereditary system in state form and use the work of A. BENSOUSSAN [1]. This leads to the study of the dual optimal control problem. It allows us to obtain the existence of the covariance operator $\Pi(t)$ and to study its properties without deriving the Riccati differential operator equation. One major difficulty is to make sense of that equation without adding any extra hypotheses on the matrices defining the original systems. For instance R. VINTER [1] and MITTER-VINTER [1] have shown that the intersection over the time t of the domains of a certain unbounded operator $\tilde{A}(t)^*$ (cf. eq. (4.23)) is generally not dense in the product space $X \times L^2(-a,0;X)$. Also the equation for the map r (cf. section 4.2) which appears in the decoupling of the optimality system (4.6)-(4.9) (cf. Theorem 4.1) does not belong to the class of hereditary systems; it could be interpreted as a special type of transport equation. But the above mentionned equations are not required to use J.C. NÉDELEC [1]'s method (see also A. BENSOUSSAN [1]). As a result we obtain a numerical scheme to compute the covariance operator for which we have convergence proofs. A similar approach has been successfully used by M.C. DELFOUR [1],[2],[3] and [4] to numerically solve the Riccati differential equation describing the evolution of the feedback operator for the linear quadratic optimal control problem.

Through several simple examples the numerical results cast some light on the nature of the operator $\Pi(t)$. In particular delays seem to create discontinuities in the derivative of the map $\alpha \mapsto \Pi_{01}(t,\alpha)$ and the covariance operator for this problem does not seem to have the same properties as the feedback operator in M.C. DELFOUR [1],[2],[3],[4].

* This research was supported in part by a "Subvention F.C.A.C. du Ministère de l'Éducation du Québec" and by National Research Council (Canada) Grant A8730 at the Centre de Recherches Mathématiques, Université de Montréal, Montréal, Québec, Canada.

Notation. Let $\mathbb{R}$ be the field of all real numbers. Let $\mathbb{R}^n$ be the Euclidean real Hilbert space of finite dimension n ($n \geq 1$, an integer). Given two real Hilbert spaces X and Y we denote by $\mathcal{L}(X,Y)$ the real Banach space of all continuous linear maps $L:X \to Y$ endowed with the natural norm $\|L\|$. The adjoint of L in $\mathcal{L}(X,Y)$ will be denoted by $L^* \in \mathcal{L}(Y,X)$. When X=Y, we write $\mathcal{L}(X,X)$ and the identity in $\mathcal{L}(X)$ is denoted by I_X. An element L of $\mathcal{L}(X)$ is said to be self adjoint (resp. positive or ≥ 0) when $L^*=L$ (resp. for all x in X, the inner product of Lx and x in X is positive or zero).

Given F a closed convex subset of $\mathbb{R}^n$ and E a real Banach space, we denote by $\mathcal{L}^p(F;E)$ the real vector space of all m-measurable (m, the Lebesgue measure on F) maps $F \to E$ which are p-integrable ($1 \leq p < \infty$) or essentially bounded ($p=\infty$). We denote by $L^p(F;E)$ the natural real Banach space associated with $\mathcal{L}^p(F;E)$ and by $\| \ \|_p$ its natural norm. When F is an interval we use the notation $L^p(a,b;E)$ where a and b are the end points of the interval. Given t_0 in $\mathbb{R}$ and $t_0 < t_1 \leq +\infty$, we denote by $H^1(t_0,t_1;E)$ the Sobolev space of all maps x in $L^2(t_0,t_1;E)$ with a distributional derivative Dx in $L^2(t_0,t_1;E)$. Let $I(a,b)$ denote the interval $]-\infty,+\infty[\cap [a,b]$. $C(a,b;E)$ will be the real Banach space of all bounded continuous maps $I(a,b) \to E$ endowed with the natural sup-norm. Given two reals $t_0 < t_1$ in $\mathbb{R}$ we define

$$\mathcal{P}(t_0,t_1) = \{(t,s) \in [t_0,t_1] \times [t_0,t_1] : t \geq s\}.$$

Finally we shall denote by $L^2_{loc}(0,\infty;E)$ and $H^1_{loc}(0,\infty;E)$ the Fréchet spaces of all maps $[0,\infty[\to E$, the restriction of which to each compact interval of the form $[0,T]$ belongs to $L^2(0,T;E)$ and $H^1(0,T;E)$.

2. System description and formulation of the problem.

2.1. Deterministic features.

Let $X = \mathbb{R}^n$, $U = \mathbb{R}^m$, $Z = \mathbb{R}^k$ for some positive non-zero integers n, m and k. Let $(\cdot,\cdot)$ and $|\cdot|$ denote the inner product and norm in X. Let $(\cdot,\cdot)_U$ (resp. $(\cdot,\cdot)_Z$) and $|\cdot|_U$ (resp. $|\cdot|_Z$) denote the inner product and norm in U (resp. Z). We are given an integer $N \geq 1$, real numbers $0 < a < +\infty$, $0 < T < +\infty$, $-a = \theta_N < \ldots < \theta_{i+1} < \theta_i < \ldots < \theta_0 = 0$. The product space $H = X \times L^2(-a,0;X)$ is endowed with inner product and norm

$$(2.1) \qquad (\!(h,k)\!) = (h^0,k^0) + \int_{-a}^{0} (h^1(\theta),k^1(\theta))\,d\theta, \quad \|h\| = (\!(h,h)\!)^{\frac{1}{2}}.$$

Let $A_i:[0,\infty[\to \mathcal{L}(X)$ $(i=0,\ldots,N)$ and $B:[0,\infty[\to \mathcal{L}(U,X)$ be measurable and bounded maps on compact intervals. Let $A_{01}:[0,\infty[\times [-a,0] \to \mathcal{L}(X)$ be also a measurable and bounded map on sets of the form $[0,t] \times [-a,0]$ for each $t \geq 0$. For f in $L^2_{loc}(0,\infty,X)$ and ξ in $L^2_{loc}(0,\infty;U)$, we consider the hereditary differential system (HDS)

$$(2.2) \qquad \begin{cases} \dfrac{dx}{dt}(t) = \displaystyle\sum_{i=0}^{N} A_i(t) \begin{Bmatrix} x(t+\theta_i) \ , & t+\theta_i \geq 0 \\ h^1(t+\theta_i), & \text{otherwise} \end{Bmatrix} + \int_{-a}^{0} A_{01}(t,\theta) \begin{Bmatrix} x(t+\theta) \ , & t+\theta \geq 0 \\ h^1(t+\theta), & \text{otherwise} \end{Bmatrix} d\theta \\ \qquad + B(t)\xi(t) + f(t) \quad \text{in} \quad]0,T[\\ x(0) = h^0, \ h = (h^0,h^1) \in H. \end{cases}$$

It is easily shown that for given h, ξ and f equation (2.2) has a unique solution x in $H^1_{loc}(0,\infty;X)$ and that the map $(h,\xi) \mapsto x: H \times L^2_{loc}(0,\infty;U) \to H^1_{loc}(0,\infty;X)$ is affine and continuous. When h belongs to the subspace

$$V = \{(h(0),h) : h \in H^1(-a,0;X)\} \tag{2.3}$$

of H, equation (2.1) can be made equivalent to the following partial differential equation (PDE) by introducing the function $y(t,\theta) = x(t+\theta)$ if $t+\theta \geq 0$ and $h(t+\theta)$ if $t+\theta < 0$:

PDE $\quad \frac{\partial}{\partial t}y(t,\theta) = \frac{\partial}{\partial \theta}y(t,\theta) \quad \text{in} \quad]0,\infty[\times]-a,0[$

BOUNDARY CONDITION $\quad \frac{\partial}{\partial t}y(t,0) = \sum_{i=0}^{N} A_i(t)y(t,\theta_i) + B(t)\xi(t) + \int_{-a}^{0} A_{01}(t,\theta)y(t,\theta)d\theta + f(t) \quad \text{in} \quad]0,\infty[$

INITIAL CONDITION $\quad y(0,\theta) = h(\theta) \quad \text{in} \quad [-a,0].$

The above formulation very naturally leads to an operational differential equation . To see this we define the state of system (2.2) as an element $\tilde{x}(t)$ of H,

$$\tilde{x}(t)^0 = x(t),\ \tilde{x}(t)^1(\theta) = \begin{Bmatrix} x(t+\theta)\ , & -t<\theta\leq 0 \\ h^1(t+\theta), & -a\leq\theta\leq -t \end{Bmatrix}. \tag{2.4}$$

We introduce the continuous linear operators $\tilde{A}_0(t):V \to X$, $\tilde{A}_1(t):V \to L^2(-a,0;X)$, $\tilde{A}(t):V \to H$, $\tilde{B}(t):U \to H$ and the vector $\tilde{f}(t)$ in H:

$$\tilde{A}_0(t)h = \sum_{i=0}^{N} A_i(t)h(\theta_i) + \int_{-a}^{0} A_{01}(t,\theta)h(\theta)d\theta,\ (\tilde{A}_1(t)h)(\theta) = \frac{dh}{d\theta}(\theta) \tag{2.5}$$

$$\tilde{A}(t)h = (\tilde{A}_0(t)h, \tilde{A}_1(t)h),\ \tilde{B}(t)w = (B(t)w,0),\ \tilde{f}(t) = (f(t),0). \tag{2.6}$$

The proofs of the following two theorems can be found in M.C. DELFOUR [5].

Theorem 2.1. For a given T>0 and all h in V, f in $L^2_{loc}(0,\infty;X)$ and ξ in $L^2_{loc}(0,\infty;U)$, $\tilde{x}$ is the unique solution in $W(0,T) = \{z \in L^2(0,T;V) : Dz \in L^2(0,T;H)\}$ (Dz denotes the distributional derivative of z) of the equation

$$\frac{dz}{dt}(t) = \tilde{A}(t)z(t) + \tilde{B}(t)\xi(t) + \tilde{f}(t) \quad \text{in }]0,T[,\ z(0) = h, \tag{2.7}$$

and the map $(h,\xi) \mapsto \tilde{x}: V \times L^2(0,T;X) \to W(0,T)$ is affine and continuous (V is endowed with the H^1-topology and W(0,T) with the norm $\|z\|_{W(0,T)} = [\|z\|^2_{L^2(0,T;V)} + \|Dz\|^2_{L^2(0,T;H)}]^{\frac{1}{2}}$). By density there exists a lifting of this map to a continuous affine map $H \times L^2(0,T;U) \to C(0,T;H)$. Moreover there exists an evolution operator $\tilde{\Phi}:\mathcal{P}(0,T) \to \mathcal{L}(H)$ such that (i) $\forall\, h \in H$, $(t,s) \mapsto \tilde{\Phi}(t,s)h$ is continuous, (ii) $\forall\, (r,s)$, $0\leq r\leq s\leq t\leq T$, $\tilde{\Phi}(t,r) = \tilde{\Phi}(t,s)\tilde{\Phi}(s,r)$, and (iii) the solution of (2.7) can be written

$$\tilde{x}(t) = \tilde{\Phi}(t,0)h + \int_0^t \tilde{\Phi}(t,r)[\tilde{B}(r)\xi(r) + \tilde{f}(r)]dr. \tag{2.8}$$

We shall also need the adjoint state. Let H' (resp. V') be the topological dual of H (resp. V). We identify elements of H and H' and denote by $\Lambda:V \to H$ and $\Lambda^*:H \to V'$

the continuous dense injections as in the theory of operational differential equations (cf. LIONS-MAGENES [1] and J.L. LIONS [1]).

Theorem 2.2. Given k in H and g in $L^2(0,T;H)$ the equations

$$\frac{dz}{dt}(t) + \tilde{A}(t)^*z(t) + \Lambda^*g(t) = 0 \ \text{ in } \ ]0,T[,\ z(T) = k, \tag{2.9}$$

has a unique solution $p(\cdot;k,g)$ in the space

$$W^*(0,T) = \{z \in C(0,T;H) : Dz \in L^2(0,T;V')\}. \tag{2.10}$$

The map $(k,g) \mapsto p(\cdot;k,g) : H \times L^2(0,T;H) \to W^*(0,T)$ is linear and continuous ($W^*(0,T)$ is endowed with the norm $\|z\|_{W^*(0,T)} = \|z\|_{C(0,T;H)} + \|Dz\|_{L^2(0,T;V')}$). Moreover

$$p(t;k,g) = \tilde{\Phi}(T,t)^*k + \int_t^T \tilde{\Phi}(r,t)^*g(r)\,dr. \quad \square \tag{2.11}$$

2.2. Stochastic features.

We now consider a noisy initial condition, that is

$$x(0) = h^0 + \zeta^0,\ x(\theta) = h^1(\theta) + \zeta^1(\theta),\ -a \le \theta < 0, \tag{2.12}$$

where $\zeta = (\zeta^0,\zeta^1)$ belong to H. From now on ξ and ζ will be the noise at the input and the noise in the initial condition, respectively. We shall also assume an observation of the form

$$z(t) = C(t)x(t) + \eta(t), \tag{2.13}$$

where $C:[0,T] \to \mathcal{L}(X,Z)$ is measurable and bounded on compact intervals and η represents the error in measurement. As in A. BENSOUSSAN [1] $\{\zeta^0,\zeta^1,\xi,\eta\}$ will be modelled as a Gaussian linear random functional on the Hilbert space $\Phi = H \times L^2(0,T;U) \times L^2(0,T;Z)$ with zero mean and covariance operator

$$\Xi = \begin{bmatrix} P_0 & 0 & 0 & 0 \\ 0 & P_1(\theta) & 0 & 0 \\ 0 & 0 & Q(t) & 0 \\ 0 & 0 & 0 & R(t) \end{bmatrix}. \tag{2.14}$$

It will be convenient to introduce the covariance operator P in $\mathcal{L}(H)$ defined as follows

$$(\!(Ph,\bar{h})\!) = (P_0h^0,\bar{h}^0) + \int_{-a}^{0} (P_1(\theta)h^1(\theta),\bar{h}^1(\theta))\,d\theta. \tag{2.15}$$

In view of the continuity of the map $(h,\xi) \mapsto x$ and the properties of the image of a linear random functional under an affine continuous map, we can look at $x(t)$ as a Gaussian linear random functional on X (for any t), where the mean of $x(t)$, $\bar{x}(t)$, is a solution of (2.2) with $\xi = 0$ and $\zeta = 0$. But it is easy to check that the mean of $\tilde{x}(t)$, $\bar{\tilde{x}}(t)$, is obtained from the mean of $x(t)$ and the mean, $\bar{h}^1(\theta)$, of $h^1(\theta)$ as follows:

$$\bar{\tilde{x}}(t)^0 = \bar{x}(t),\ \bar{\tilde{x}}(t)^1(\theta) = \begin{Bmatrix} \bar{x}(t+\theta)\ ,\ t+\theta \ge 0 \\ \bar{h}^1(t+\theta),\ \text{otherwise} \end{Bmatrix}. \tag{2.16}$$

As a result $\bar{\tilde{x}}(t)$ is a solution of the state equation (2.7) with $\xi = 0$ and $\zeta = 0$ and the covariance operator $\Gamma(t)$ of $\tilde{x}(t)$ is a "weak solution" of the equation

(2.17) $$\frac{d\Gamma}{dt}(t) = \tilde{A}(t)\Gamma(t) + \Gamma(t)\tilde{A}(t)^* + \tilde{B}(t)Q(t)\tilde{B}(t)^*,\ \Gamma(0) = P.$$

2.3. Formulation of the problem.

For each T we want to determine the best estimator of the linear random functional $\tilde{x}(T)$ with respect to the linear random functional $z(s)$, $0 \le s \le T$. It is a linear random functional $\hat{\tilde{x}}(T)$ which can be obtained (see A. BENSOUSSAN [1]) through the following control problem. We start with the deterministic system (2.7) with the initial condition (2.17) at time 0 and consider the variables ξ and $\zeta = (\zeta^0,\zeta^1)$ as control variables. We want to minimize the cost function for a given h.

(2.18) $$J_T(\xi,\zeta,h) = ((P^{-1}\zeta,\zeta)) + \int_0^T (Q(t)^{-1}\xi(t),\xi(t))_U dt + \int_0^T (R(t)^{-1}(z(t)-C(t)x(t),z(t)-C(t)x(t))_Z dt$$

(provided that P, Q(t) and R(t) be invertible almost everywhere).

3. Solution of the optimal control problem (2.7)-(2.17)-(2.18).

It will be technically advantageous to work in the state space. For this purpose we redefine the cost function (2.18) in terms of the state

(3.1) $$\frac{d\tilde{x}}{dt}(t) = \tilde{A}(t)\tilde{x}(t) + \tilde{B}(t)\xi(t) + \tilde{f}(t) \quad \text{in }]0,T[,\ \tilde{x}(0) = h + \zeta$$

(3.2) $$J_T(\xi,\zeta) = ((P^{-1}\zeta,\zeta)) + \int_0^T (Q(t)^{-1}\xi(t),\xi(t))_U dt + \int_0^T (R(t)^{-1}(z(t)-\tilde{C}(t)\tilde{x}(t)),z(t)-\tilde{C}(t)\tilde{x}(t))_Z dt$$

where $\tilde{C}(t)h = C(t)h^0$. It will be convenient to introduce the variables y and $\bar{y}$:

(3.3) $$\frac{dy}{dt}(t) = \tilde{A}(t)y(t) + \tilde{B}(t)\xi(t) \quad \text{in }]0,T[,\ y(0) = \zeta,$$

(3.4) $$\frac{d\bar{y}}{dt}(t) = \tilde{A}(t)\bar{y}(t) + \tilde{f}(t) \quad \text{in }]0,T[,\ \bar{y}(0) = h.$$

We notice that $\tilde{x}(t) = y(t) + \bar{y}(t)$ and rewrite the cost function (3.2) in terms of y

(3.5) $$J_T(\xi,\zeta) = ((P^{-1}\zeta,\zeta)) + \int_0^T (Q(t)^{-1}\xi(t),\xi(t))_U dt + \int_0^T (R(t)^{-1}\tilde{C}(t)y(t),\tilde{C}(t)y(t))_Z dt$$
$$-2\int_0^T (R(t)^{-1}[z(t)-\tilde{C}(t)\bar{y}(t)],\tilde{C}(t)y(t))_Z dt + \text{terms independent of } \xi \text{ and } \zeta.$$

The pair $(\hat{\xi},\hat{\zeta})$ which minimizes the cost function over all (ξ,ζ) in $L^2(0,T;U) \times H$ is characterized by

(3.6) $$((P^{-1}\hat{\zeta},\zeta)) + \int_0^T (Q(t)^{-1}\hat{\xi}(t),\xi(t))_U dt + \int_0^T (R(t)^{-1}\tilde{C}(t)\hat{y}(t),\tilde{C}(t)y(t))_Z dt$$
$$= \int_0^T (\tilde{C}(t)\hat{y}(t),R(t)^{-1}[z(t)-\tilde{C}(t)\bar{y}(t)])_Z dt \quad \forall\ \xi,\zeta.$$

By introducing the adjoint system

$$(3.7)\quad \begin{cases} \frac{d\hat{p}}{dt}(t) + \tilde{A}(t)^*\hat{p}(t) + \Lambda^*\tilde{C}(t)^*R(t)^{-1}[\tilde{C}(t)\hat{y}(t)-(z(t)-\tilde{C}(t)\bar{y}(t))] = 0 \quad \text{in} \quad]0,T[\\ \hat{p}(T) = 0, \end{cases}$$

we obtain that $(\hat{\xi},\hat{\zeta})$ are characterized by

$$(3.8)\qquad \hat{\zeta} = -P\hat{p}(0), \quad \hat{\xi}(t) = -Q(t)\tilde{B}(t)^*\hat{p}(t).$$

By substituting $\hat{\zeta}$ and $\hat{\xi}$ in (3.3) we obtain

$$(3.9)\qquad \frac{d\hat{y}}{dt}(t) = \tilde{A}(t)\hat{y}(t) - \tilde{B}(t)Q(t)\tilde{B}(t)^*\hat{p}(t) \quad \text{in} \quad]0,T[, \ \hat{y}(0) = -P\hat{p}(0),$$

and system (3.9)-(3.8) is called the <u>optimality system</u>. The optimal $\hat{\hat{x}}$ corresponding to $(\hat{\zeta},\hat{\xi})$ is given by $\hat{\hat{x}}(t) = \hat{y}(t) + \bar{y}(t)$.

4. <u>Dual optimal control problem.</u>

We introduce the maps $g:[0,T] \to H$ and $N:[0,T] \to \mathcal{L}(H)$,

$$(4.1)\qquad g(t) = \tilde{C}(t)^*R(t)^{-1}[\tilde{C}(t)\bar{y}(t)-z(t)], \ N(t) = \tilde{B}(t)Q(t)\tilde{B}(t)^*,$$

and we consider the following control system and its associated cost function

$$(4.2)\qquad \frac{dp}{dt}(t) + \tilde{A}(t)^*p(t) + \Lambda^*[\tilde{C}(t)^*w(t)+g(t)] = 0 \quad \text{in} \quad]0,T[, \ p(T) = k$$

$$(4.3)\qquad J^*(w,k) = ((Pp(0),p(0))) + \int_0^T [((N(t)p(t),p(t)))+(R(t)w(t),w(t))_Z]dt,$$

where k and w belong to H and $L^2(0,T;Z)$, respectively. Given k there exists a unique $\hat{w}$ in $L^2(0,T;Z)$ which minimizes $J^*(w,k)$ over all w in $L^2(0,T;Z)$. The minimizing $\hat{w}$ is characterized by Euler's equation:

$$(4.4)\qquad ((P\hat{p}(0),q(0))) + \int_0^T [((N(t)\hat{p}(t),q(t)))+(R(t)\hat{w}(t),w(t))_Z]dt = 0, \ \forall\, w,$$

where

$$(4.5)\qquad \frac{dq}{dt}(t) + \tilde{A}(t)^*q(t) + \Lambda^*\tilde{C}(t)^*w(t) = 0 \quad \text{in} \quad]0,T[, \ q(T) = 0.$$

If we introduce the dual system of system (4.2)

$$(4.6)^*\qquad \frac{d\hat{y}}{dt}(t) = \tilde{A}(t)\hat{y}(t) - N(t)\hat{p}(t) \quad \text{in} \quad]0,T[, \ \hat{y}(0) = -P\hat{p}(0),$$

identity (4.4) reduces to

$$(4.7)\qquad \forall\, m \in L^2(0,T;Z), \ \int_0^T (R(t)\hat{w}(t)-\tilde{C}(t)\hat{y}(t),w(t))_Z dt = 0.$$

Since R(t) is invertible (4.7) is equivalent to

$$(4.8)\qquad \hat{w}(t) = R(t)^{-1}\tilde{C}(t)\hat{y}(t), \ \text{a.e.} \quad \text{in} \quad [0,T].$$

If we now substitute identity (4.8) into equation (4.2) we obtain

$$(4.9)\qquad \frac{d\hat{p}}{dt}(t) + \tilde{A}(t)^*\hat{p}(t) + \Lambda^*[\tilde{C}(t)^*R(t)^{-1}\tilde{C}(t)\hat{y}(t)+g(t)] = 0 \quad \text{in} \quad]0,T[, \ \hat{p}(T) = k.$$

---

* This equation must be interpreted in an appropriate weak sense, namely,

$$\hat{y}(t) = -\tilde{\Phi}(t,0)P\hat{p}(0) - \int_0^t \tilde{\Phi}(t,r)N(r)\hat{p}(r)dr \quad \text{in} \quad [0,T].$$

Equations (4.6)-(4.9) form the optimality system of problem (4.2)-(4.3). The reader will notice that system (4.6)-(4.9) with k=0 is identical to the optimality system (3.7)-(3.9) with the significant difference that (4.6)-(4.9) was derived without the hypothesis that P and Q(t) be invertible.

Remark. In the formulation of the dual optimal control problem $\tilde{C}$ can be any strongly measurable map $[0,T] \to \mathcal{L}(H,Z)$ which is bounded in [0,T]. It is not necessary to restrict ourselves to $\tilde{C}$'s of the form $\tilde{C}(t)h = C(t)h^0$ for some $C:[0,T] \to \mathcal{L}(X,Z)$.

We shall now proceed to the decoupling of optimality system (4.6)-(4.9) and to the study of the decoupling operator as in J.L. LIONS [2] and A. BENSOUSSAN [1].

In the sequel we shall use the notation $M(t) = \tilde{C}(t)^* R(t)^{-1}\tilde{C}(t)$.

Theorem 4.1. Let $\hat{p}$ and $\hat{y}$ be the solution of system (4.6)-(4.9). Then there exists a family of linear operators $\Pi(t): H \to H$ and a family of elements r(t) in H, $0 \le t \le T$, such that

(4.10) $$\hat{y}(t) = -\Pi(t)\hat{p}(t) + r(t), \quad \text{in} \quad [0,T].$$

$\Pi(t)$ and r(t) are obtained in the following manner: (i) we solve the system

(4.11) $$\begin{cases} \frac{d\beta}{dt}(t) = \tilde{A}(t)\beta(t) - N(t)\gamma(t) \quad \text{in} \quad]0,s[,\ \beta(0) = -P\gamma(0) \\ \frac{d\gamma}{dt}(t) + \tilde{A}(t)^*\gamma(t) + \Lambda^* M(t)\beta(t) = 0 \quad \text{in} \quad]0,s[,\ \gamma(s) = k \end{cases}$$

and $\Pi(s)k = -\beta(s)$; (ii) we solve the system

(4.12) $$\begin{cases} \frac{d\eta}{dt}(t) = \tilde{A}(t)\eta(t) - N(t)\chi(t) \quad \text{in} \quad]0,s[,\ \eta(0) = -P\chi(0) \\ \frac{d\chi}{dt}(t) + \tilde{A}(t)^*\chi(t) + \Lambda^*[M(t)\eta(t)+g(t)] = 0 \quad \text{in} \quad]0,s[,\ \chi(s) = 0 \end{cases}$$

and $r(s) = \eta(s)$. □

4.1. Study of the operator $\Pi(t)$.

In order to study the operator $\Pi(t)$ we make use of Theorem 4.1 and consider system (4.2) with g=0 and the cost function (4.3) in a time interval [0,s] for some s in]0,T].

Theorem 4.2. (i) If we denote by γ (resp. $\bar{\gamma}$) the solution of equation

(4.13) $$\frac{d\gamma}{dt}(t) + [\tilde{A}(t)^* - \Lambda^* M(t)\Pi(t)]\gamma(t) = 0 \quad \text{in} \quad]0,s[,\ \gamma(s) = k \ (\text{resp. } \bar{k}),$$

then

(4.14) $$(\!(\Pi(s)k,\bar{k})\!) = (\!(P\gamma(0),\bar{\gamma}(0))\!) + \int_0^s (\!([N(r)+\Pi(r)M(r)\Pi(r)]\gamma(r),\bar{\gamma}(r))\!)dr$$

and in particular if $\hat{w}$ is the optimal control corresponding to k

(4.15) $$J_s^*(\hat{w},k) = (\!(\Pi(s)k,k)\!).$$

(ii) The operator $\Pi(s)$ is a self adjoint element of $\mathcal{L}(H)$, there exists a constant $c > 0$ (independent of s and h) such that

(4.16) $$\forall\ s, \quad \forall\ k, \quad \|\Pi(s)k\| \le c\|k\|,$$

and the map $s \mapsto \Pi(s): [0,T] \to \mathcal{L}(H)$ is weakly continuous (hence strongly measurable and bounded). $\Pi(s)$ can be decomposed in a unique way into a matrix of operators

$$\begin{bmatrix}\Pi_{00}(s), & \Pi_{01}(s)\\ \Pi_{10}(s), & \Pi_{11}(s)\end{bmatrix} \quad \begin{matrix}\Pi_{00}(s) \in \mathcal{L}(X),\ \Pi_{01}(s) \in \mathcal{L}(L^2(-a,0;X),X)\\ \Pi_{10}(s) \in \mathcal{L}(X,L^2(-a,0;X)),\ \Pi_{11}(s) \in \mathcal{L}(L^2(-a,0;X)).\end{matrix} \tag{4.17}$$

Moreover

$$\Pi_{00}(s)^* = \Pi_{00}(s) \geq 0,\ \Pi_{01}(s) = \Pi_{10}(s)^*,\ \Pi_{11}(s)^* = \Pi_{11}(s) \geq 0, \tag{4.18}$$

and

$$J_s^*(\hat{m},k) = (\Pi_{00}(s),k^0,k^0) + 2(\Pi_{01}(s)k^1,k^0) + (\Pi_{11}(s)k^1,k^1)_2. \tag{4.19}$$

(iii) The equation

$$y_s(t) = \tilde{\Phi}(t,s)h - \int_s^t \tilde{\Phi}(t,r)\Pi(r)M(r)y_s(r)dr \quad \text{in} \quad [s,T] \tag{4.20}$$

has a unique solution y_s in $C(s,T;H)$ which generates an evolution operator $\Lambda(t,s)$ defined as $\Lambda(t,s)h = y_s(t)$ with the following properties:

a) $\forall\ (t,s) \in \wp(0,T),\ \Lambda(t,s) \in \mathcal{L}(H)$;

b) $\forall\ 0 \leq s \leq r \leq t \leq T,\ \Lambda(t,s) = \Lambda(t,r)\Lambda(r,s)$;

c) $\forall\ h \in H,\ (t,s) \mapsto \Lambda(t,s)h:\wp(0,T) \to H$ is continuous.

Equation (4.14) can now be rewritten in the form

$$(\!(\Pi(s)k,\bar{k})\!) = (\!(P\Lambda(s,0)^*k,\Lambda(s,0)^*\bar{k})\!)+\int_0^s (\!([N(r)+\Pi(r)M(r)\Pi(r)]\Lambda(s,r)^*k,\Lambda(s,r)\bar{k})\!)dr. \ \square \tag{4.21}$$

Equation (4.21) is the "integral form" of the desired Riccati operator differential equation. It is identical to the one in MITTER-VINTER [1]. Formally one should obtain in $]0,T[$ an equation of the form

$$\frac{d\Pi}{dt}(t) = \Pi(t)\tilde{A}(t)^* + \tilde{A}(t)\Pi(t) - \Pi(t)M(t)\Pi(t) + N(t),\ \Pi(0) = P. \tag{4.22}$$

However it is not easy to interpret this equation properly since $\tilde{A}(t)$ and $\tilde{A}(t)^*$ are unbounded operators which depend on time. In particular

$$\bigcap_{0\leq t\leq T} \{k \in H:\tilde{A}(t)^*k \in H\} \tag{4.23}$$

is not necessarily dense in H (cf. R. VINTER [1] and MITTER-VINTER [1]).

Remark. Notice that the equation

$$\frac{dp}{dt}(t) + \tilde{A}(t)^*p(t) - \Lambda^*M(t)\Pi(t)p(t) = 0 \quad \text{in} \quad]0,T[,\ p(T) = k$$

is perfectly legitimate and that its solution is $p(t) = \Lambda(T,t)^*k$.

4.2. Study of the function $r(t)$.

In order to study the function r we consider the problem (4.2)-(4.3) with $k=0$. By Theorem 4.1 we know that

$$r(t) = \hat{y}(t) + \Pi(t)\hat{p}(t). \tag{4.24}$$

Straightforward computations using equations (4.6), (4.9) and (4.21) will show that

$$\begin{cases}\dfrac{dr}{dt}(t) = [\tilde{A}(t)-\Pi(t)M(t)\Lambda]r(t) - \Pi(t)g(t), \quad \text{in} \quad]0,T[,\\ r(0) = 0,\end{cases} \tag{4.25}$$

where the above equation is to be interpreted as

$$(4.26)\qquad r(t) = -\int_0^t \Lambda(t,s)\Pi(s)g(s)ds \quad \text{in} \quad [0,T]$$

or

$$(4.27)\qquad r(t) = -\int_0^T \tilde{\Phi}(t,s)\Pi(s)[M(s)r(s)+g(s)]ds \quad \text{in} \quad [0,T].$$

5. Approximation of the dual system.

In this section we shall exploit earlier results of M. DELFOUR [1],[4] on the standard optimal control problem. We assume that we can find non zero positive integers $M, L, L_0, \ldots, L_N$ and a discretization step $\delta > 0$ such that $T = M\delta$, $a = L\delta$, $\theta_i = -L_i\delta$, $i=0,\ldots,N$.

5.1. Approximation of initial data.

We approximate the product space $H = X \times L^2(-b,0;X)$ by the finite dimensional space $H^\delta = X^{L+1}$ endowed with inner products

$$(5.1)\qquad (\underline{h},\underline{k})_\delta = (h_0,k_0) + \delta \sum_{\ell=-L}^{-1} (h_\ell,k_\ell), (\underline{h},\underline{k})_L = \sum_{\ell=-L}^{0} (h_\ell,k_\ell).$$

We introduce the maps

$$(5.2)\qquad h = (h^0,h^1) \mapsto r_\delta(h) = (h^0,h^1_{-1},\ldots,h^1_{-L}):H \to H^\delta,$$

$$(5.3)\qquad \underline{h} = (h_0,h_{-1},\ldots,h_{-L}) \mapsto q_\delta(\underline{h}) = (h_0, \sum_{\ell=-L}^{-1} h_\ell \chi_\ell):H^\delta \to H,$$

where χ_ℓ is the characteristic function of $[\ell\delta,(\ell+1)\delta[$ and

$$(5.4)\qquad h^1_\ell = \frac{1}{\delta}\int_{\ell\delta}^{(\ell+1)\delta} h^1(\theta)d\theta, \; -L \le \ell \le -1.$$

It is readily seen that $\|q_\delta r_\delta(h)\| \le \|h\|$. We shall also need the transformation

$$i_\delta:H^\delta \to H^\delta, \; [i_\delta(\underline{h})]_0 = h_0, \; [i_\delta(\underline{h})]_\ell = \delta^{\frac{1}{2}} h_\ell, \; \ell=-L,\ldots,-1.$$

If we introduce the map $I_\delta = i_\delta^2$, we notice that for all $\underline{h}$ and $\underline{k}$

$$(\underline{h},\underline{k})_\delta = (i_\delta\underline{h},i_\delta\underline{k})_L = (i_\delta^2\underline{h},\underline{k})_L = (I_\delta\underline{h},\underline{k})_L.$$

5.2. Approximation of the differential equation.

We associate with A_i a family of matrices and with f an element $\underline{f}$ in X^M

$$(5.5)\qquad A_i^m = \frac{1}{\delta}\int_{m\delta}^{(m+1)\delta} A_i(t)dt, \; (i=0,\ldots,N), \; f^m = \frac{1}{\delta}\int_{m\delta}^{(m+1)\delta} f(t)dt, \; m=0,\ldots,M-1.$$

With A_{01} we associate the family of matrices

$$(5.6)\qquad \begin{cases} A_{01}^{m,0} = \dfrac{1}{\delta^2}\displaystyle\int_{m\delta}^{(m+1)\delta} dt \int_{-(t-m\delta)}^{0} d\theta\, A_{01}(t,\theta), A_{01}^{m,-L} = \dfrac{1}{\delta^2}\displaystyle\int_{m\delta}^{(m+1)\delta} dt \int_{-b}^{-a+(m+1)\delta-t} d\theta\, A_{01}(t,\theta) \\ A_{01}^{m,\ell} = \dfrac{1}{\delta^2}\displaystyle\int_{m\delta}^{(m+1)\delta} dt \int_{(m+\ell)\delta-t}^{(m+\ell+1)\delta-t} d\theta\, A_{01}(t,\theta), \; \ell=-1,\ldots,-(L-1), \end{cases}$$

We construct a finite dimensional approximation $U^\delta = U^M$ to the space of control maps $L^2(0,T;U)$. We define the maps

$$\xi \mapsto r_\delta(\xi) = (\xi_0,\dots,\xi_{M-1}):L^2(0,T;U) \to U^M, \tag{5.7}$$

$$\xi_m = \frac{1}{\delta}\int_{m\delta}^{(m+1)\delta} \xi(t)\,dt,\ m=0,\dots,M-1, \tag{5.8}$$

$$\xi = (\xi_0,\dots,\xi_{M-1}) \mapsto q_\delta(\xi) = \sum_{m=0}^{M-1} \xi_m\chi_m : U^M \to L^2(0,T;U), \tag{5.9}$$

where χ_m is the characteristic function of $[m\delta,(m+1)\delta[$. It is reasily seen that $\|q_\delta r_\delta(\xi)\|_2 \le \|\xi\|_2$. It will be clear from the context whether q_δ and r_δ are associated with $L^2(0,T;U)$ or with the product space $H = X \times L^2(-b,0;X)$. Finally we associate with B the following family of matrices

$$B^m = \frac{1}{\delta}\int_{m\delta}^{(m+1)\delta} B(t)\,dt,\ m=0,\dots,M-1. \tag{5.10}$$

We now associate with $\underline{h}$ in H^δ, $\underline{\xi} = (\xi_0,\dots,\xi_{M-1})$ and $\underline{f} = (f_0,\dots,f_{M-1})$ the following numerical scheme

$$\begin{cases} x_{m+1}-x_m=\delta\Big[\sum_{i=0}^{N} A_i^m \begin{Bmatrix} x_{m-L_i}, & m-L_i\ge 0 \\ h_{m-L_i}, & m-L_i<0\end{Bmatrix} + \sum_{\ell=-L}^{0} \delta\, A_{01}^{m,\ell} \begin{Bmatrix} x_{m+\ell}, & m+\ell\ge 0 \\ h_{m+\ell}, & m+\ell<0\end{Bmatrix} + B^m\xi_m + f^m\Big], \\ \qquad\qquad m=0,\dots,M-1, \\ x_0 = h_0. \end{cases} \tag{5.11}$$

The following propositions summarize the results we shall need.

Proposition 5.1. (Stability). We denote by $(x_0^\delta,\dots,x_{M-1}^\delta)$ the unique solution of (5.11) corresponding to $\underline{h} = r_\delta(h)$, $\underline{\xi} = r_\delta(\xi)$ and $\underline{f} = r_\delta(f)$ for h in H, ξ in $L^2(0,T;U)$ and f in $L^2(0,T;X)$. We define the maps

$$x^\delta(t) = \sum_{m=0}^{M-1} x_m^\delta\chi_m(t),\ x^\delta(T) = x_M^\delta,\ \ Dx^\delta(t) = \sum_{m=0}^{M-1} \frac{x_{m+1}^\delta - x_m^\delta}{\delta}\chi_m(t), \tag{5.12}$$

where χ_m denotes the characteristic function of $[m\delta,(m+1)\delta[$. As δ goes to zero there exists a constant $c > 0$ (independent of h, f, v and δ) such that

$$\max\{|x_m^\delta|:m=0,\dots,M\} + \|x^\delta\|_\infty + \|Dx^\delta\|_2 \le c[\|h\|+\|f\|_2+\|\xi\|_2]. \quad \square \tag{5.13}$$

Proposition 5.2. (Convergence). Fix h in H, f in $L^2(0,T;X)$, ξ in $L^2(0,T;U)$. As δ goes to zero with $M\delta = T$

$$\max\{|x_m^\delta - x(m\delta)|:0\le m\le M\} + \|x^\delta - x\|_2 + \|Dx^\delta - Dx\|_2 \tag{5.14}$$

converges to zero, where x is the solution in $W^{1,2}(0,T;X)$ of equation (2.2). $\square$

Corollary. Assume that $A_1,\dots,A_N$ and B are constant matrices and that A_{01} is identically zero. As δ goes to zero there exists a constant $c > 0$ (independent of δ, h, v and f) such that

$$\|x-x^{\delta}\|_2 + \max\{|x(m\delta)-x_m^{\delta}|:0\le m\le M\} \le c\delta\|Dx\|_2. \quad \square \tag{5.15}$$

5.3. Approximation of the differential equation for the state.

We now introduce an explicit finite difference scheme to approximate the differential equation for the state. Given $\underline{h}$ in H^{δ}, we want to determine $\{x_{m,n}:0\le m\le M, -L\le n\le 0\}$ from the following set of equations

$$x_{m+1,0}-x_{m,0} = \delta\Big[\sum_{i=0}^{N} A_i^m x_{m,-L_i} + \delta\sum_{\ell=-L}^{0} A_{01}^{m,\ell} x_{m,\ell} + B^m \xi_m + f^m\Big].\ m=0,\dots,M-1, \tag{5.16}$$

$$x_{0,\ell} = h_\ell,\ \ell = -L,\dots,0,\quad x_{m+1,\ell} = x_{m,\ell+1},\quad m=0,\dots,M-1,\quad \ell=-L,\dots,-1. \tag{5.17}$$

We define an $(L+1)n \times (L+1)n$ matrix $\tilde{A}^m$ and an $(L+1)n \times m$ matrix $\tilde{B}^m$

$$[\tilde{A}^m\underline{h}]_0 = \sum_{i=0}^{N} A_i^m h_{-L_i} + \delta\sum_{\ell=-L}^{0} A_{01}^{m,\ell} h_\ell,\ [\tilde{A}^m\underline{h}]_\ell = \frac{1}{\delta}[h_{\ell+1}-h_\ell],\ \ell=-L,\dots,-1,$$

$$[\tilde{B}^m w]_0 = B^m w,\quad [\tilde{B}^m w]_\ell = 0,\quad \ell=-L,\dots,-1,$$

and the vectors $\tilde{x}_m$ and $\tilde{f}^m$ in H^{δ}

$$\tilde{x}_m = (x_{m,0},x_{m,-1},\dots,x_{m,-L}),\quad \tilde{f}^m = (f^m,0,\dots,0). \tag{5.18}$$

Equations (5.16) and (5.17) can now be rewritten in the more compact form

$$\tilde{x}_{m+1} - \tilde{x}_m = \delta[\tilde{A}^m\tilde{x}_m + \tilde{B}^m\xi_m + \tilde{f}^m],\quad m=0,\dots,M-1,\quad \tilde{x}_0 = \underline{h}. \tag{5.19}$$

We notice that equation (5.19) has been constructed in such a way that $x_{m,\ell}$ remains constant along the characteristics of the differential equation for the state

$$x_{m+1,\ell} = x_{m,\ell+1},\quad m=0,\dots,M-1,\quad \ell=-L,\dots,-1.$$

As a result there exists $\underline{y} = (y_0,\dots,y_M)$ such that $x_{m,\ell} = y_{m+\ell}$, $m+\ell \ge 0$ and it is easy to see that $(y_0,\dots,y_M)$ is the solution of (5.11) with initial condition $\underline{h}$. Hence the scheme (5.19) has a unique solution.

Proposition 5.4. Let $(\tilde{x}_0,\dots,\tilde{x}_M)$ be the solution of scheme (5.19) for $\underline{h} = r_\delta(h)$, $\underline{\xi} = r_\delta(\xi)$ and $\underline{f} = r_\delta(f)$ for some h in H, ξ in $L^2(0,T;U)$ and f in $L^2(0,T;X)$. We define the map $\tilde{x}^{\delta}:[0,T] \to H$ as follows

$$\begin{cases} \tilde{x}^{\delta}(t)^0 = \sum_{m=0}^{M-1} x_{m,0}\chi_m(t), t \in [0,T[,\ \tilde{x}^{\delta}(T) = q_\delta(\tilde{x}_M), \\ \tilde{x}^{\delta}(t)^1(\theta) = \sum_{m=0}^{M-1}\sum_{n=-L}^{-1} [x_{m,n}\chi_{m,n}^{\ell}(t,\theta) + x_{m,n+1}\chi_{m,n}^{u}(t,\theta)],\ (t,\theta) \in [0,T[\times [-a,0], \end{cases} \tag{5.20}$$

where χ_m is the characteristic function of $[m\delta,(m+1)\delta[$, $\chi_{m,n}^{\ell}$ is the characteristic function of

$$\{(t,\theta) \in [m\delta,(m+1)\delta[\times [n\delta,(n+1)\delta[: t + \theta < (m+n+1)\delta\}, \tag{5.21}$$

and $\chi_{m,n}^{u}$ the characteristic function of

$$\{(t,\theta) \in [m\delta,(m+1)\delta[\times [n\delta,(n+1)\delta[: (m+n+1)\delta \le t+\theta\}. \tag{5.22}$$

(i) (Stability). There exists a constant $c > 0$, independent of δ, h, ξ and f, such that for all h in H, f in $L^2(0,T;X)$ and ξ in $L^2(0,T;U)$

(5.23) $$\max\{\|q_\delta(\tilde{x}_m)\|:m=0,\dots,M\} + \|\tilde{x}^\delta\|_\infty \le c[\|h\|+\|f\|_2+\|\xi\|_2].$$

(ii) (Convergence). Fix h, f and ξ. As δ goes to zero with $M\delta = T$

(5.24) $$\max\{\|q_\delta(\tilde{x}_m)-\tilde{x}(m\delta)\|:m=0,\dots,M\} \to 0,$$

and $\tilde{x}^\delta$ converges to $\tilde{x}$ in $L^\infty(0,T;H)$, where $\tilde{x}$ denotes the solution of equation (2.7).□

5.4. Approximation of the differential equation for the adjoint state.

In this section we introduce an approximation of the adjoint state equation (2.9). Consider the following scheme

(5.25) $$p_{m+1} - p_m + \delta[I_\delta^{-1}(\tilde{A}^m)^* I_\delta p_{m+1}+g^m] = 0, \quad m=0,\dots,M-1, \text{ in } H^\delta, \quad p_M = \underline{k} \text{ in } H^\delta$$

where $\underline{k} = r_\delta(k)$ and $\underline{g} = (g^0,\dots,g^{M-1})$ is constructed from g in $L^2(0,T;H)$

(5.26) $$g^m = \frac{1}{\delta}\int_{m\delta}^{(m+1)\delta} r_\delta(g(t))dt, \quad m=0,\dots,M-1.$$

Proposition 5.5. Let $(p_0,\dots,p_M)$ be the solution of the scheme (5.25) with final datum $\underline{k}$ for some k in H and $\underline{g} = (g^0,g^1,\dots,g^{M-1})$ for some g in $L^2(0,T;H)$. Let the approximation $p^\delta:[0,T] \to H$ be defined as

(5.27) $$p^\delta(t) = \sum_{m=0}^{M-1} q_\delta(p_{m+1})\chi_m(t), \quad 0 \le t < T, \quad p^\delta(T) = q_\delta(p_M).$$

(i) (Stability). There exists a constant $c > 0$ (independent of δ, h and g) such that

(5.28) $$\max\{\|q_\delta(p_m)\|:m=0,\dots,M\} + \|p^\delta\|_\infty \le c[\|k\|+\|g\|_2].$$

(ii) (Convergence). Let p denote the solution of equation (2.9) in the space $W^*(0,T)$. Then for all h in H

$$\max\{|((q_\delta(p_m),h))-((p(m\delta),h))|:m=0,\dots,M\} \to 0 \text{ as } \delta \to 0 \text{ with } M\delta = T$$

and the map $t \mapsto ((h,p^\delta(t)))$ converges to the map $t \mapsto ((h,p(t)))$ in $L^\infty(0,T;\mathbb{R})$. □

Corollary. Assume that the matrices $A_1,\dots,A_N$ are constant, that A_{01} is identically zero and that $k = (k^0,0)$ and $g(t) = (g^0(t),0)$. As δ goes to zero there exists a constant $c' > 0$ (independent of k^0, g^0 and δ) such that

(5.29) $$\|p-p^\delta\|_2 + \max\{\|p(m\delta)-q_\delta(p_m)\|:0\le m\le M\} \le c'\delta[|k^0|+\|g^0\|_2]. \quad □$$

6. Approximation of the dual optimal control problem.

We now construct an approximation to the dual optimal control problem of section 2.4. We start with the approximation (5.25) to system (4.2):

(6.1) $$\begin{cases} p_{m+1}-p_m+\delta[I_\delta^{-1}(\tilde{A}^m)^* I_\delta p_{m+1}+(\tilde{\underline{C}}^m)^* w_m+g^m] = 0, & m=0,\dots,M-1, \text{ in } H^\delta \\ p_M=\underline{k}=r_\delta(k) \text{ for some k in H}, & \underline{w}=(w_0,\dots,w_{M-1})=r_\delta(w) \text{ for } w \in L^2(0,T;Z), \end{cases}$$

where

$$(6.2)\qquad g^m = \frac{1}{\delta}\int_{m\delta}^{(m+1)\delta} r_\delta[g(t)]dt,\ \underline{\tilde{C}}^m = \frac{1}{\delta}\int_{m\delta}^{(m+1)\delta} \tilde{C}(t)q_\delta dt,\ m = 0,\ldots,M-1.$$

We associate with system (6.1) the following approximation of the cost function

$$(6.3)\qquad J_\delta^*(\underline{w},\underline{k}) = (\underline{P}p_0,p_0)_L + \delta\sum_{m=0}^{M-1}[(\underline{N}^m p_{m+1},p_{m+1})_L + (\underline{R}^m w_m,w_m)_Z],$$

where

$$(6.4)\qquad \left\{\begin{array}{l} \underline{P} = q_\delta^* P q_\delta,\ \underline{N}^m = \underline{\tilde{B}}^m \underline{Q}^m(\underline{\tilde{B}}^m)^*,\ \underline{R}^m = \frac{1}{\delta}\int_{m\delta}^{(m+1)\delta} R(t)dt \\ \underline{Q}^m = \frac{1}{\delta}\int_{m\delta}^{(m+1)\delta} Q(t)dt,\ \underline{\tilde{B}}^m = \frac{1}{\delta}\int_{m\delta}^{(m+1)\delta} r_\delta(\tilde{B}(t))dt. \end{array}\right\},\ m = 0,\ldots,M-1,$$

The approximate optimal control problem consists in minimizing $J_\delta^*(\underline{w},\underline{k})$ over all $\underline{w}$ in Z^δ:

$$\mathrm{Inf}\{J_\delta^*(\underline{w},\underline{k}):\underline{w} \in Z^\delta\}.$$

Lemma 6.1. Given $\underline{k}$ in H^δ, the approximate optimal control problem (6.1)-(6.3) has a unique solution $\underline{\hat{w}}$ in U^δ. This solution is completely characterized by the optimality system:

$$(6.5)\qquad \hat{p}_{m+1}-\hat{p}_m+\delta[I_\delta^{-1}(\tilde{A}^m)^* I_\delta \hat{p}_{m+1}+(\tilde{C}^m)^*\hat{w}_m+g^m] = 0,\ 0 \le m \le M-1,\ \hat{p}_M = \underline{k},$$

$$(6.6)\qquad \hat{y}_{m+1}-\hat{y}_m = \delta[\tilde{A}^m\hat{y}_m - I_\delta^{-1}\underline{N}^m\hat{p}_{m+1}],\ 0 \le m \le M-1,\ \hat{y}_0 = I_\delta^{-1}\underline{P}\hat{p}_0,$$

$$(6.7)\qquad \hat{w}_m = (\underline{R}^m)^{-1}\underline{\tilde{C}}^m I_\delta \hat{y}_m,\quad 0 \le m \le M-1.\ \square$$

Proposition 6.2. (i) Given h in H, for each $\delta > 0$ the approximate optimal control problem with initial condition $\underline{h} = r_\delta(h)$ has a unique solution $\underline{\hat{w}}$ in U^δ. As δ goes to zero, $q_\delta(\underline{\hat{w}})$ converges to $\hat{w}$ in $L^2(0,T;Z)$, where $\hat{w}$ is the optimal control in the minimization problem (4.2)-(4.3). (ii) We define the maps p^δ and $y^\delta:[0,T] \to H$

$$(6.8)\qquad p^\delta(t) = \sum_{m=0}^{M-1} q_\delta(\hat{p}_m)\chi_m(t),\ 0 \le t < T,\ p^\delta(T) = q_\delta(\hat{p}_M),$$

$$(6.9)\qquad y^\delta(t) = \sum_{m=0}^{M-1} q_\delta(\hat{y}_m)\chi_m(t),\ 0 \le t < T,\ y^\delta(T) = q_\delta(\hat{y}_M).$$

As δ goes to zero $J_\delta(\underline{\hat{w}},r_\delta(h))$ converges to $J(\hat{w},h)$,

$$(6.10)\qquad \left\{\begin{array}{l} \forall\ s \in [0,T],\ q_\delta(\hat{p}_m) \to \hat{p}(s) \text{ in H weak with } m\delta = s \\ \forall\ h \in H, \text{ the map } t \mapsto ((p^\delta(t),h)) \text{ converges to} \\ \qquad \text{the map } t \mapsto ((\hat{p}(t),h)) \text{ in } L^\infty(0,T;\mathbb{R}), \end{array}\right.$$

$$(6.11)\qquad \left\{\begin{array}{l} \forall\ s \in [0,T],\ q_\delta(\hat{y}_m) \to \hat{y}(s) \text{ in H weak with } m\delta = s \\ \forall\ h \in H, \text{ the map } t \mapsto ((y^\delta(t),h)) \text{ converges to} \\ \qquad \text{the map } t \mapsto ((\hat{y}(t),h)) \text{ in } L^\infty(0,T;\mathbb{R}), \end{array}\right.$$

where $\hat{p}$ and $\hat{y}$ are the solutions of system (4.6)-(4.9). □

Proposition 6.3. Let the sequences $\{\hat{p}_m\}$ and $\{\hat{y}_m\}$ be the solutions of the optimality equations (6.5) to (6.7). There exists a family of matrices $\{\Pi_m^\delta : m = 0,\dots,M\}$ in $\mathcal{L}(H^\delta)$ and a family of elements $\{\rho_m : m = 0,\dots,M\}$ in H^δ such that

$$(6.12)\qquad I_\delta \hat{y}_m = -\Pi_m^\delta \hat{p}_m + \rho_m,\quad m = 0,\dots,M.$$

Moreover Π_m^δ and ρ_m are obtained in the following manner: (i) we solve the system

$$(6.13)\qquad \gamma_{m+1}-\gamma_m+\delta[I_\delta^{-1}(\tilde{A}^m)^* I_\delta \gamma_{m+1}+\underline{M}^m I_\delta \beta_m],\ 0 \le m \le r-1,\ \gamma_r = \underline{k},$$

$$(6.14)\qquad \beta_{m+1}-\beta_m = \delta[\tilde{A}^m \beta_m - I_\delta^{-1}\underline{N}^m \gamma_{m+1}],\ 0 \le m \le r-1,\ \beta_0 = I_\delta^{-1}\underline{P}\gamma_0,$$

and $\Pi_r^\delta \underline{k} = -I_\delta \beta_r$ (where $\underline{M}^m = (\tilde{\underline{C}}^m)^*(\underline{R}^m)^{-1}\tilde{\underline{C}}^m$); (ii) we solve the system

$$(6.15)\qquad \xi_{m+1}-\xi_m+\delta[I_\delta^{-1}(\tilde{A}^m)^* I_\delta \xi_{m+1}+\underline{M}^m I_\delta \eta_m + g^m] = 0,\ 0 \le m \le r-1,\ \xi_r = \underline{k},$$

$$(6.16)\qquad \eta_{m+1}-\eta_m = \delta[\tilde{A}^m \eta_m - I_\delta^{-1}\underline{N}^m \xi_{m+1}],\ 0 \le m \le r-1,\ \eta_0 = I_\delta^{-1}\underline{P}\xi_0$$

and $\rho_r = I_\delta \eta_r$. □

7. Approximation of the equations for Π and ρ.

To study the family of operators $\Pi(s)$, $0 < s \le T$, we have considered the optimal control problem in the interval $[0,s]$ with $g=0$. By analogy we fix an integer r, $0 < r \le M$, and consider the system

$$(7.1)\qquad p_{m+1}-p_m+\delta[I_\delta^{-1}(\tilde{A}^m)^* I_\delta p_{m+1}+(\tilde{C}^m)^* w_m] = 0,\ m = 0,\dots,M-1,\ p_M = \underline{k},$$

and the optimal control problem for the cost function $J^*_{\delta,r}(\underline{w},\underline{k})$ (that is, (6.3) on $[0,r]$) associated with (7.1).

Proposition 7.1. We fix an integer $r > 0$. (i) If $\hat{\underline{w}}$ is the minimizing control

$$(7.2)\qquad J^*_{\delta,r}(\hat{\underline{w}},\underline{k}) = (\Pi_r^\delta \underline{k},\underline{k})_L,\ \Pi_r^\delta = (\Pi_r^\delta)^* \ge 0,$$

where * denotes the adjoint in H^δ with respect to the inner product $(\ ,\ )_L$. (ii) As δ goes to zero, there exists a constant $c > 0$ such that

$$(7.3)\qquad \forall\, k \in H,\ (\Pi_r^\delta r_\delta(k), r_\delta(k))_L \le c\|k\|^2.$$

(iii) If γ_m, β_m, $m = 0,\dots,M$, are the solutions of (6.13)-(6.14), then $\Pi_m^\delta \gamma_m = -I_\delta \beta_m$, $r \le m \le M$. □

Theorem 7.2. We define

$$(7.4)\qquad P_m = r_\delta^* \Pi_m^\delta r_\delta,\ 0 \le m \le M.$$

We introduce the map $\Pi^\delta : [0,T] \to \mathcal{L}(H)$

$$(7.5)\qquad \Pi^\delta(t) = \sum_{m=0}^{M-1} P_m \chi_m(t)\ \text{in}\ [0,T[,\ \Pi^\delta(T) = P_M.$$

(i) (Stability). As δ goes to zero, there exists a constant $c > 0$ (independent of δ) such that

(7.6) $$\max\{\|P_m\|_{\mathcal{L}(H)} : m = 0,\dots,M\} \le c.$$

(ii) (Convergence). For fixed s in [0,T] and all h and k in H

(7.7) $$((P_m h,k)) \to ((\Pi(s)h,k)) \text{ as } \delta \to 0 \text{ with } m\delta = s. \quad \square$$

Corollary. We can easily verify that

(7.8) $$[P_m]_{00} = [\Pi_m^\delta]_{00},\ [P_m]_{01}h^1 = \int_{-a}^{0} \sum_{n=-L}^{-1} \delta^{-1}[\Pi_m^\delta]_{0n}\chi_n(\theta)h^1(\theta)d\theta$$

(7.9) $$([P_m]_{11}h^1)(\alpha) = \int_{-a}^{0} \sum_{\ell=-L}^{-1} \sum_{n=-L}^{-1} \delta^{-2}[\Pi_m^\delta]_{\ell n}\chi_\ell(\alpha)\chi_n(\theta)h^1(\theta)d\theta.$$

For fixed s in [0,T], as δ goes to zero with $m\delta = s$

(7.10) $$[P_m]_{00} \to [\Pi(s)]_{00} \text{ in } \mathcal{L}(X),\ \forall\, h^1 \in L^2(-a,0;X),\ [P_m]_{01}h^1 \to [\Pi(s)]_{01}h^1 \text{ in } X,$$

(7.11) $$\forall\, h^1 \in L^2(-a,0;X),\ [P_m]_{11}h^1 \to [\Pi(s)]_{11}h^1 \text{ in } L^2(-a,0;X) \text{ weak}$$

and the norms of $[P_m]_{00}$, $[P_m]_{01}$ and $[P_m]_{11}$ are uniformly bounded. $\square$

Theorem 7.3. For δ small enough the family Π_m^δ, $0 \le m \le M$, as defined by equations (6.13)-(6.14) of Proposition 6.3 is the solution of the following set of equations:

(7.12) $$\begin{cases} \Pi_{m+1} = \delta \underline{N}^m + (I+\delta I_\delta \tilde{A}^m I_\delta^{-1})\Pi_m(I+\delta \underline{M}^m \Pi_m)^{-1}(I+\delta I_\delta \tilde{A}^m I_\delta^{-1})^*,\ 0 < m \le M, \\ \Pi_0 = \underline{P}, \text{ where I is the identity matrix in } \mathcal{L}(H^\delta). \quad \square \end{cases}$$

To study the family of vectors $\rho(s)$, $0 < s \le T$, we have considered the optimal control problem in the interval [0,s] with k=0. By analogy we can fix an integer r, $0 < r \le M$, and consider the system (6.1) with $\underline{k}=0$ and the optimal control problem for the associated cost function (6.3). We will obtain a set of equations for ρ_m.

8. Numerical examples.

In this section we consider a number of examples which will illustrate the behaviour of the map $\alpha \mapsto \Pi_{01}(t,\alpha)$ as a function of the time t in [0,T]. In all examples a=1, T=2, $X = U = Z = \mathbb{R}$ and the observation equation is $z(t) = x(t)+\eta(t)$ with $R(t) = 1$, $0 \le t \le 2$. All equations are to be interpreted in an appropriate way as stochastic differential equations.

Example 1. This example has an analytical solution. Consider

(8.1) $$\dot{x}(t) = x(t-1),\ 0 \le t \le 2,\ x(0) = h^0+\zeta^0,\ x(\theta) = h^1(\theta),\ -1 \le \theta < 0.$$

Let $c = P_0^{-1}$. It can be shown that

(8.2) $$\Pi_{00}(s) = \begin{Bmatrix} (s+c)^{-1}, & 0 \le s \le 1 \\ s^2(1+c+\frac{s^3-1}{3})^{-1}, & 1 \le s \le 2 \end{Bmatrix}$$

$$\Pi_{01}(s,\theta) = \begin{Bmatrix} (s+c)^{-1}\chi_{[-s,0]}(\theta) & , 0\le s\le 1 \\ s(1+c+\frac{s^3-1}{3})^{-1}\begin{Bmatrix} s+\theta, & 1-s\le\theta\le 0 \\ 1, & \text{otherwise}\end{Bmatrix}, & 1\le s\le 2\end{Bmatrix}. \tag{8.3}$$

The results appear in Figures 1a and 1b for $P_0 = 1$.

Example 2. We now consider a system with two delays

$$\dot{x}(t) = x(t-\tfrac{1}{2})+x(t-1),\ 0\le t\le 2,\ x(0) = h^0+\zeta^0,\ x(\theta) = h^1(\theta),\ -1\le\theta<0. \tag{8.4}$$

The results appear in Figure 2 for $P_0 = 1$.

Example 3. We allow input noise in (8.4)

$$\dot{x}(t) = x(t-\tfrac{1}{2})+x(t-1)+\xi(t),\ 0\le t\le 2,\ x(0) = h^0+\zeta^0,\ x(\theta) = h^1(\theta),\ -1\le\theta<0. \tag{8.5}$$

The results appear in Figure 3 for $P_0 = 1$ and $Q(t) = 1$, $0\le t\le 2$.

Example 4. We consider system (8.5) but without noise in the initial condition, $x(0) = h^0$. The results are shown in Figure 4.

Example 5. We consider the system

$$\dot{x}(t) = A_1(t)x(t-1),\ 0\le t\le 2,\ x(0) = h^0+\zeta^0,\ x(\theta) = h^1(\theta),\ -1\le\theta<0. \tag{8.6}$$

The results appear in Figures 1a and 5 for $P_0 = 1$ and

$$A_1(t) = \begin{Bmatrix} 1, & n/10 \le t < (n+1)/10, & n \text{ even} \\ 0, & n/10 \le t < (n+1)/10, & n \text{ odd}\end{Bmatrix}. \tag{8.7}$$

Example 6. It is similar to Example 5 but without noise in the initial condition and with input noise:

$$\dot{x}(t) = A_1(t)x(t-1)+\xi(t),\ 0\le t\le 2,\ x(0) = h^0,\ x(\theta) = h^1(\theta),\ -1\le\theta<0. \tag{8.8}$$

The results appear in Figure 6 for $Q(t) = 1$, $0\le t\le 2$, and A_1 as in (8.7).

References.

A. BENSOUSSAN [1], Filtrage optimal des systèmes linéaires, Dunod, Paris, 1971.
[2], Filtrage des systèmes linéaires avec retard, IRIA Report INF 7118/71027; Oct. 1971.

A. BENSOUSSAN, M.C. DELFOUR and S.K. MITTER [1], Optimal filtering for linear stochastic hereditary differential systems, Proc. 1972 IEEE Conference on Decision and Control, New Orleans, 378-380.

R.T. CURTAIN [1], A Kalman-Bucy theory for affine hereditary differential equations, in Control theory, numerical methods and computer systems modelling, eds. A. Bensoussan and J.L. Lions, Springer-Verlag, New York 1975, 22-43.

M.C. DELFOUR [1], Solution numérique de l'équation différentielle de Riccati rencontrée en théorie de la commande optimale des systèmes héréditaires linéaires, in Control theory, numerical methods and computer systems modelling, eds. A. Bensoussan and J.L. Lions, Springer-Verlag, New York 1975, 362-383.
[2], Numerical solution of the optimal control problem for linear hereditary differential systems with a linear-quadratic cost function and approximation of the Riccati differential equation, CRM-Report 408, Université de Montréal, Montréal, Canada 1974.
[3], Numerical solution of the operational Riccati differential equation in the optimal control theory of linear hereditary differential systems with a linear-quadratic cost function, Proc. 1974 IEEE Conference on Decision and Control, Phoenix, 784-790.

[4], The linear quadratic optimal control problem for hereditary differential systems: theory and numerical solution, to appear in Journal of Applied Mathematics and Optimization.
[5], State theory of linear hereditary differential systems, Centre de Recherches Mathématiques Report CRM-395, Université de Montréal, Montréal, Canada.

H. KWAKERNAAK [1], Optimal filtering in linear systems with time delay, IEEE Trans. Automatic Control 12 (1967), 169-173.

R. KWONG [1], Doctoral dissertation, Massachusetts Institute of Technology, 1975.

A. LINDQUIST [1], A theorem on duality between estimation and control for linear stochastic system with time delay, J. Math. Anal. Appl. 37 (1972), 516-536.

J.L. LIONS [1], Equations différentielles opérationnelles, Springer-Verlag, Berlin 1961.
[2], Optimal control of systems governed by partial differential equations, Springer-Verlag, New York, 1971.

J.L. LIONS and E. MAGENES [1], Problèmes aux limites non homogènes, Vol. 1,2,3, Paris, Dunod 1968.

S.K. MITTER and R. VINTER [1], Filtering for linear stochastic hereditary differential systems, in Control theory, numerical methods and computer systems modelling, eds. A. Bensoussan and J.L. Lions, Springer-Verlag, New York 1975, 1-21.

J.C. NÉDELEC [1], Schémas d'approximation pour des équations intégro-différentielles de Riccati, Thèse de doctorat d'état, Paris, 1970.

M. SENDAULA [1], On the duality between estimation and control for linear hereditary differential systems, Internal Report, Makerere University, Kampala, Uganda.

R. VINTER [1], On the evolution of the state of linear differential delay equations in M^2: properties of the generator, Report ESL-R-541, Electronic Systems Laboratory, M.I.T., Cambridge, Mass., U.S.A..

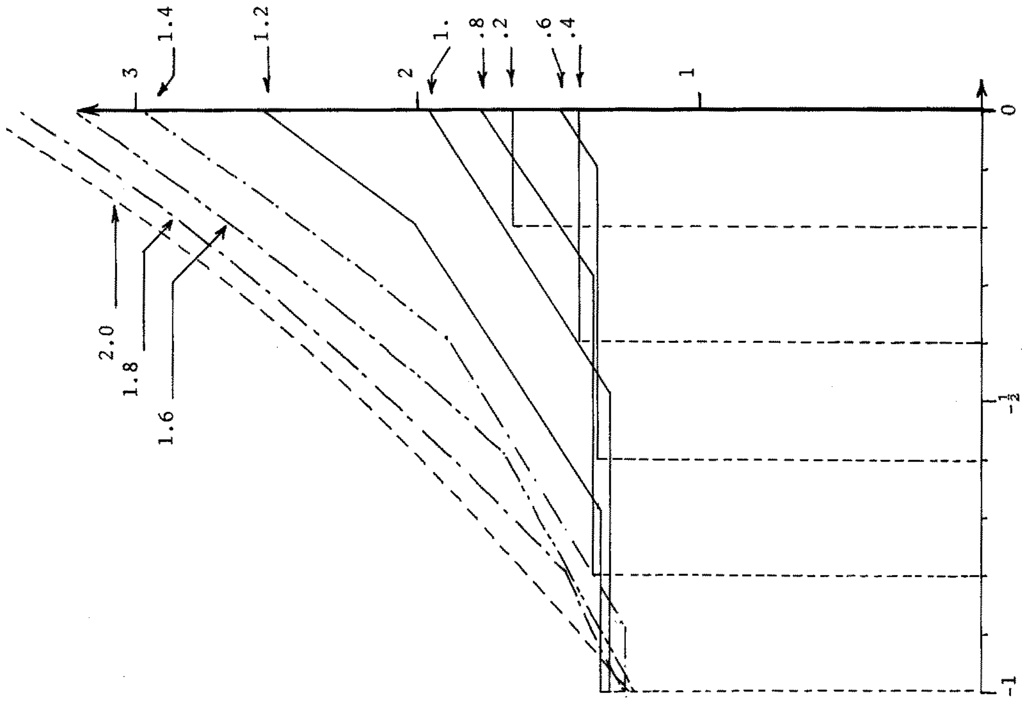

Figure 2.

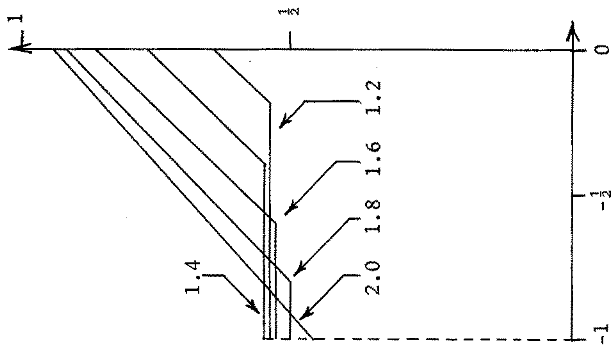

Figure 1b.

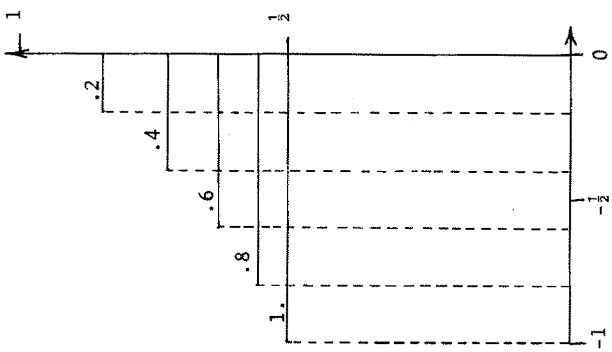

Figure 1a.

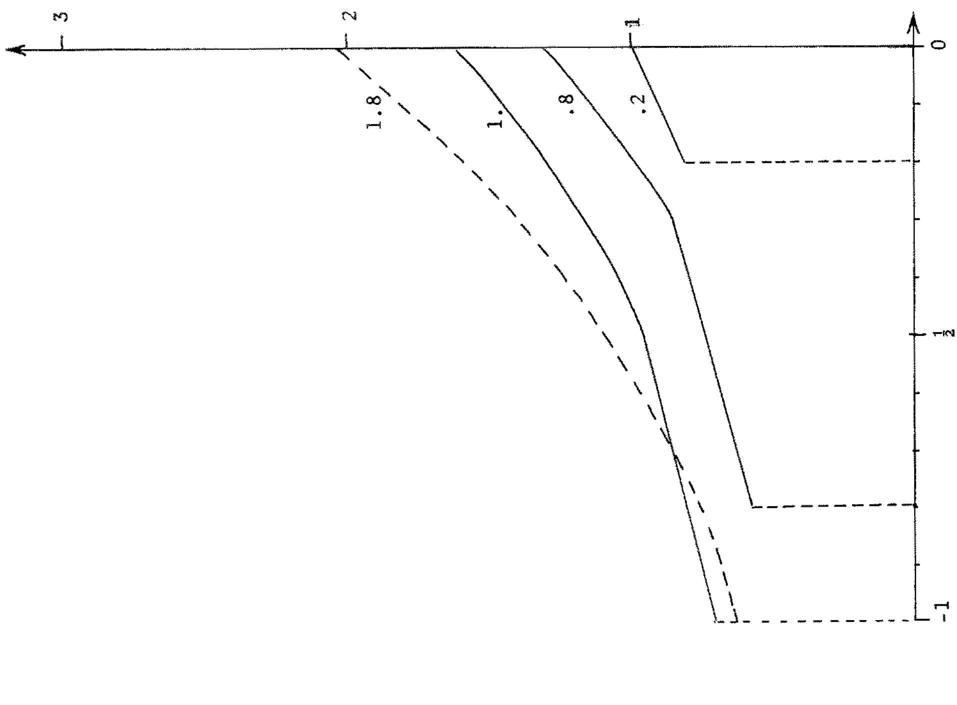

Figure 3.

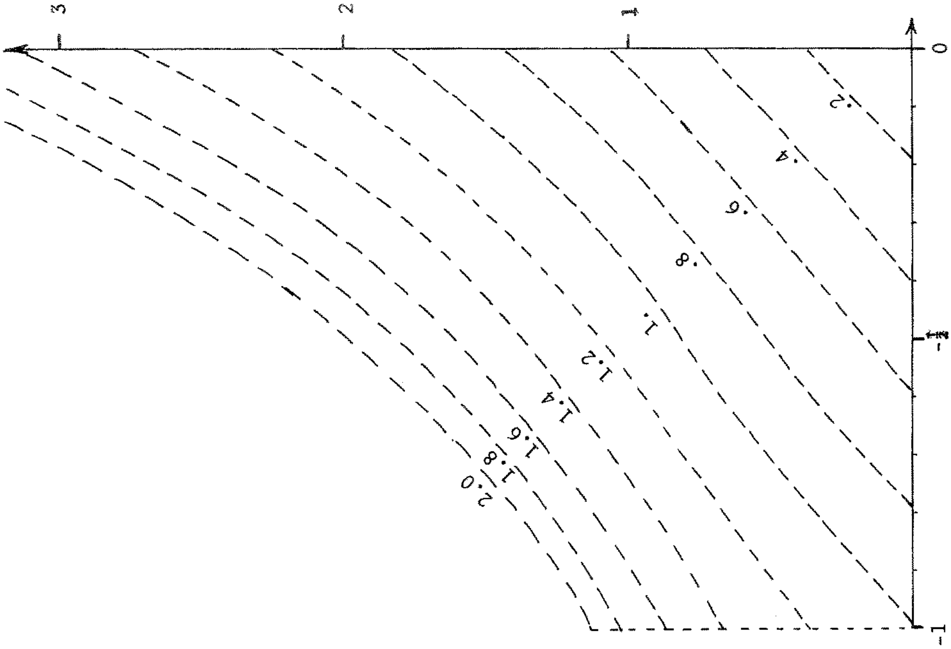

Figure 4.

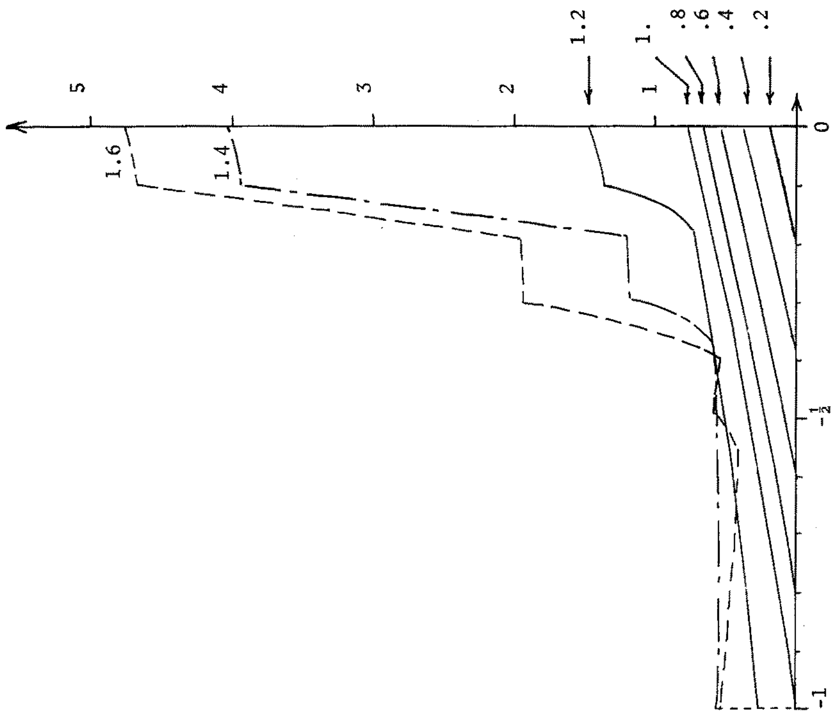

Figure 6.

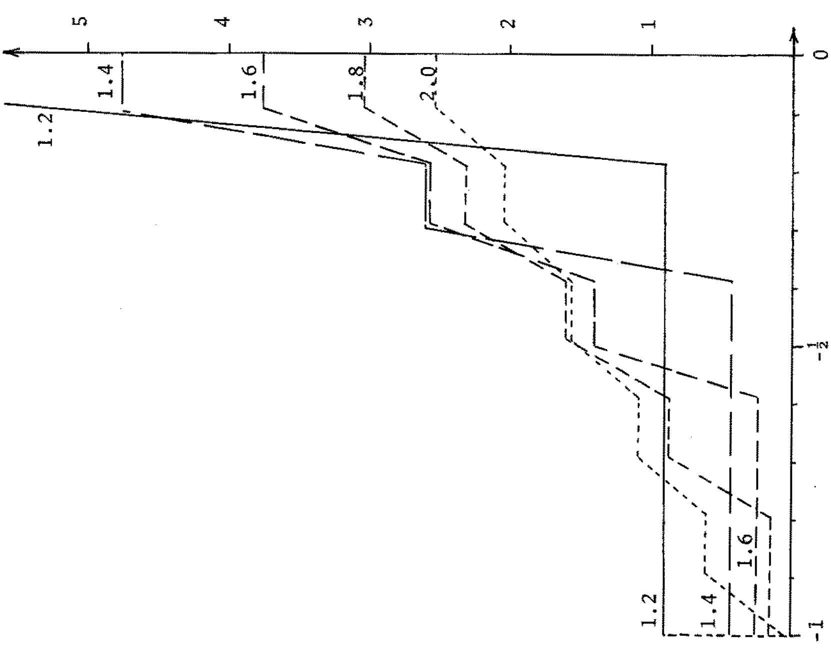

Figure 5.