# Expected values, feared values, and partial information optimal control

Pierre Bernhard

january, 1995

**Abstract**

We show how a morphism between the ordinary algebra $(+, \times)$ and the $(\max, +)$ algebra offer a completely paprallel treatment of stochastic and minimax control of disturbed nonlinear systems with partial information.

## 1 Introduction

Minimax control, or worst case design, as a means of dealing with uncertainty is an old idea. It has gained a new popularity with the recognition, in 1988, of the fact that $H_\infty$-optimal control could be cast into that concept. Although some work in that direction existed long before (see [8]), this viewpoint has vastly renewed the topic. See [3] and related work.

Many have tried to extend this work to a nonlinear setup. Most prominent among them perhaps is the work of Isidori, [15] citeisiast92 but many others have followed suit : [23] [4] [5] and more recently [18] [19]. This has contributed to a renewed interest in nonlinear minimax control.

We insist that the viewpoint taken here is squarely that of minimax control, *and not nonlinear $H_\infty$-optimal control*. Several reasons for that claim. Fot one thing, we only consider finite time problems, and therefore do not consider stability issues which are usually central in $H_\infty$-optimal control. We dont stress quadratic loss functions. But more importantly, we claim that the minimax problem is only an intermediary step in $H_\infty$ theory, used to insure existence of a fixed point to the feedback equations $z = P_K w$, $w = \Delta P z$ ($P_K$ is the controlled plant, $\Delta P$ the model uncertainty). In that respect, the nonlinear equivalent is *not* the minimax problem usually considered, but rather the contraction problem independently tackeled by [14].

If we decide that minimax is an alternative to stochastic treatment of disturbances (input uncertainties, rather than plant uncertainties), it makes sense to try to establish a parallel. In this direction, we have the striking morphism developed by Quadrat and coworkers, see [21] [2] [1]. We shall review here recent work, mainly by ourselves, Baras, and James, in the light of this parallel, or Quadrat's morphism. This paper is in a large extent based upon [7].

## 2 Quadrat's morphism

In a series of papers [21] [2] [1], giving credit to other authors for early developements, Quadrat and coauthors have fully taken advantage of the morphism introduced between the ordinary algebra $(+, \times)$ and the $(\min, +)$, or alternatively the $(\max +)$, algebra to develop a *decision calculus* parallel to probability calculus. It has been pointed out by Quadrat and coauthors that a possible way of understanding that morphism was through Cramer's transform. We shall not, however, develop that way of thinking here, but merely rely on the algebraic similarity between the two calculus.

Let us briefly review some concepts, based on [1].

### 2.1 Cost measure

The parallel to a probability measure is a *cost measure*. Let $\Omega$ be a topological space, $\mathcal{A}$ a $\sigma$-field of subsets, $K : \mathcal{A} \to \mathbb{R} \cup \{-\infty\}$ is called a cost measure if it satisfies the following axioms :

- $K(\emptyset) = -\infty$
- $K(\Omega) = 0$
- for any family of (disjoint) elements $A_n$ of $\mathcal{A}$,

$$K(\cup A_n) = \sup_n K(A_n)\,.$$

(It is straightforward to see that the word "disjoint" can be omitted from this axiom).

One may notice the parallel with a probability measure. In the first two axioms, the O of probability measures, the neutral element of the addition, is replaced by the neutral element of the max operator : $-\infty$, and the 1, the neutral element of the product, is replaced by the neutral element of the sum, 0. In the third axiom, the sum of the measures of the disjoint sets is replaced by the max.

The function $\bar{G} : \Omega \to \mathbb{R} \cup \{-\infty\}$ is called a *cost density* of $K$ if we have

$$\forall A \in \mathcal{A}, \quad K(A) = \sup_{\omega \in \Omega} \bar{G}(\omega)\,.$$

One has the following theorem (Akian)

**Theorem 1** *Every cost measure defined on the open sets of a Polish space $\Omega$ admits a unique maximal extension to $2^{\Omega}$, this extension has a density, which is a concave u.s.c. function.*

## 2.2 Feared values

The wording *Feared values* is introduced here to stress the parallel with *expected values.*

When a stochastic disturbance is introduced into a problem model, in order to derive a controller design for instance, it comes with a given probability distribution. We shall always assume here that these distributions have densities. Let therefore $w \in \mathsf{W}$ be a stochastic variable. If its probability disribution is $\Pi(\cdot)$. Let $\psi$ be a function of $w$ with values in $\mathbb{R}$ We define its expected value as

$$\mathbb{E}_w \psi := \int \psi(w)\Pi(w)\,dw$$

and we omit the subscript $w$ to $\mathbb{E}$ when no ambiguity results. Similarly, let a disturbance $w$ be given together with a cost distribution $\Gamma(\cdot)$. The *feared value* of a function $\psi$ from $\mathcal{W}$ into $\mathbb{R}$ is defined as

$$\mathbb{F}_w \psi := \max_w [\psi(w) + \Gamma(w)]$$

which is the formula dual to that of the expected value in Quadrat's morphism.

The "Fear" operator enjoys the linearity properties one would exect in the (max, +) algebra :

$$\mathbb{F}(\max\{\phi, \psi\}) = \max\{\mathbb{F}\phi, \mathbb{F}\psi\}\,,$$

and if $\lambda$ is a constant,

$$\mathbb{F}(\lambda + \psi) = \lambda + \mathbb{F}\psi\,.$$

A sequence of stochastic variables $\{w_t\}$, $t = 0 \dots T-1$ also denoted $w_{[0,T]}$, are said to be *independant* if their joint probability density is the product of their individual probability densities $\Pi_t$:

$$\Pi(w_{[0,T]}) = \prod_{t=0}^{T-1} \Pi_t(w_t)$$

leading to the following formula, where $J$ is a function of the whole sequence

$$\mathbb{E}J(w_{[0,T]}) = \int J(w_{[0,T]}) \prod_{t=0}^{T} \Pi_t(w_t)\,dw_{[0,T]}\,.$$

In a similar fashion, a sequence of independant decision variables $w_{[0,T]}$ with cost densities $\Gamma_t$ will have a joint cost density $\Gamma$ equal to the sum of their individual cost densities:

$$\Gamma(w_{[0,T]}) = \sum_{t=0}^{T-1} \Gamma_t(w_t)$$

leading to the dual formula

$$\mathbb{F}J(w_{[0,T]}) = \max_{w_{[0,T]}} [J(w_{[0,T]}) + \sum_{t=0}^{T} \Gamma_t(w_t)]\,.$$

**Conditioning** Let a pair of decision variables $(v, w)$ ranging over sets $\mathsf{V} \times \mathsf{W}$ have a joint cost density $r(v, w)$. We may define the marginal law for $v$ as

$$p(v) = \max_{w \in \mathsf{W}} r(v, w)$$

for which it is true that the feared value of the characteristic function $1\!\!1_A(v)$ of a set $A \subset \mathsf{V}$ is given by

$$\mathbb{F} 1\!\!1_A = \max_{v \in \mathsf{V}} p(v)$$

preserving the duality with the probabilistic formulas

$$p(v) = \int_{\mathsf{W}} r(v, w)\, dw$$

and

$$\mathbb{E} 1\!\!1_A = \int_{\mathsf{V}} p(v)\, dv\,.$$

Similarily, we have the dual of Bayes formula, defining the *conditional cost measure* $q(w|v)$ as

$$q(w|v) = r(v, w) - p(v)$$

Let $\mathbb{F}^v_w$ denote the corresponding feared value, we have the "embeded algebra" formula:

$$\mathbb{F}_v \left[ \mathbb{F}^v_w \psi(v, w) \right] = \mathbb{F} \psi(v, w)\,.$$

We shall often need a less simple form of conditioning such as (with transparent notations)

$$\mathbb{F}[\psi(w) \mid w \in A] = \max_{w \in A} [\psi(w) + \Gamma(w)]\,.$$

which should clearly be seen as the basic conditioning operation.

# 3 The discrete time control problem

## 3.1 The problem

We consider a partially observed two input control system

$$x_{t+1} = f_t(x_t, u_t, w_t)\,, \tag{1}$$
$$y_t = h_t(x_t, w_t)\,, \tag{2}$$

where $x_t \in \mathbb{R}^n$ is the state at time $t$, $u_t \in \mathsf{U}$ the (minimizer's) control, $w_t \in \mathsf{W}$ the disturbance input, and $y_t \in \mathsf{Y}$ the measured output. We shall call $\mathbf{U}$ the set of input sequences over the time horizon $[0, T]$: $\{u_t\}_{t \in [0,T]}$ usually written as $u_{[0,T]} \in \mathbf{U}$, and likewise for $w_{[0,T]} \in \mathbf{W}$. The initial state $x_0 \in \mathsf{X}_0$ is also considered part of the disturbance. We shall call $\omega = (x_0, w_{[0,T]})$ the combined disturbance, and $\Omega = \mathsf{X}_0 \times \mathbf{W}$ the set of disturbances.

The solution of (1) (2) above shall be written as

$$\begin{aligned} x_t &= \phi_t(u_{[0,T]}, \omega)\,, \\ y_t &= \eta_t(u_{[0,T]}, \omega)\,. \end{aligned}$$

Finally, we shall call $u^t$ a partial sequence $(u_0, u_1, \ldots, u_t)$ and $\mathsf{U}^t$ the set of such sequences [1], likewise for $w^t \in \mathsf{W}^t$ and $y^t \in \mathsf{Y}^t$. Also, we write $\omega^t = (x_0, w^t) \in \Omega^t$.

The solution of (1) and (2) may alternatively be written as

$$x_t = \phi_t(u^{t-1}, \omega^{t-1})\,, \tag{3}$$

$$y_t = \eta_t(u^{t-1}, \omega^t)\,. \tag{4}$$

We shall also write

$$x^t = \phi^t(u^{t-1}, \omega^{t-1})\,, \tag{5}$$

$$y^t = \eta^t(u^{t-1}, \omega^t)\,, \tag{6}$$

to refer to the partial sequences solution of (1) and (2)

Admissible controllers will be *strictly causal output feedbacks* of the form $u_t = \mu_t(u^{t-1}, y^{t-1})$. We denote by $\mathcal{M}$ the class of such controllers.

A performance index is given. In general, it may be of the form

$$J(x_0, u_{[0,T]}, w_{[0,T]}) = M(x_T) + \sum_{t=0}^{T-1} L_t(x_t, u_t, w_t)\,.$$

However, we know that, to the expense of increasing the state dimension by one if necessary, we can always bring it back to a purely terminal payoff of the form

$$J(x_0, u_{[0,T]}, w_{[0,T]}) = M(x_T) = M \circ \phi_T(u_{[0,T]}, \omega)\,. \tag{7}$$

The data of a strategy $\mu \in \mathcal{M}$ and of a disturbance $\omega \in \Omega$ generates through (1)(2) a unique pair of sequences $(u_{[0,T]}, w_{[0,T]}) \in \mathbf{U} \times \mathbf{W}$. Thus, with no ambiguity, we may also use the abusive notation $J(\mu, \omega)$. The aim of the control is to minimize $J$, in some sense, "in spite of the unpredictable disturbances".

We want to compare here two ways of turning this unprecise statement into a meaningful mathematical problem.

In the first approach, *stochastic control*, we modelize the unknown disturbance as a random variable, more specifically here a random variable $x_0$ with a probability density $N(x)$ and an independant white stochastic process $w_{[0,T]}$ of known instantaneous probability distribution $\Pi_t$. (We notice that nothing

[1] notice the slight inconsistency in notations, in that our $\mathsf{U}^t$ is the cartesian $(t+1)$ power of $\mathsf{U}$. Other choices of notations have their drawbacks too.

in the sequel prevents $\Pi_t$ from depending on $x_t$ and $u_t$.) The criterion to be minimzed is then

$$H(\mu) := \mathbb{E}_\omega J(\mu, \omega) \,. \tag{8}$$

This can be expanded into

$$H(\mu) = \int M(x_T) \left( \prod_{t=0}^{T-1} \Pi_t(w_t) \right) N(\xi)\, dw_{[0,T]}\, d\xi$$

In the second approach, we are given the cost density $N$ of $x_0$, and the cost densities $\Gamma_t$ of the $w_t$'s. (Again, $\Gamma_t$ might depend on $x_t$ and $u_t$.) The criterion to be minimized is then

$$G(\mu) := \mathbb{F}_\omega J(\mu, \omega) \,, \tag{9}$$

which can be expanded into

$$G(\mu) := \max_\omega [M(x_T) + \sum_{t=0}^{T-1} \Gamma_t(w_t) + N(x_0)]$$

**Remark** If all cost measures of the disturbances are taken constant, (e.g. 0), then $G(\mu)$ is, if it exists, the *guaranteed value* given only the sets over which the perturbations range. Therefore, minimizing it is insuring the best possible guaranteed value.

## 3.2 Dynamic programming

### 3.2.1 Stochastic dynamic programming

We quickly recall here for reference purpouses the classical solution of the stochastic problem via dynamic programming. One has to introduce the *conditional state probability measure*, and, assuming it is absolutely continuous with respect to the Lebesgue measure, its density $W$. Let, thus, $W_t(x)\, dx$ be the conditional probability measure of $x_t$ given $y^{t-1}$, or *a priori* state probability distribution at time $t$, and $W_t^\eta(x)\, dx$ be the conditional state distribution given $y^{t-1}$ and given that $y_t = \eta$, or *a posteriori* state probability distribution at time $t$.

Clearly, $W_t$ is a function only of past measurements. As a matter of fact, we can give the filter that lets one compute it. Starting from

$$W_0(x) = N(x) \tag{10}$$

at each step, $W_t^\eta$ can be obtained by Bayes rule. A standard condition for this step to be well posed is that, for all $(t, x, w)$, the map $w \mapsto h_t(x, w)$ be locally onto, and more specifically that the partial derivative $\partial h_t(x, w)/\partial w$ be invertible. It suffices here to notice that, because the information is increasing, (the information algebras are nested), we have, for any test function $\psi(\cdot) \in L^1(\mathbb{R}^n)$,

$$\mathbb{E}_y \int \psi(x) W_t^y(x)\, dx = \int \psi(x) W_t(x)\, dx \,. \qquad (11)$$

Then $W_{t+1}$ is obtained by propagating $W_t^{y_t}$ through the dynamics. It suffices for our purpouse to define this propagation by the dual operator: for any test function $\psi$,

$$\int \psi(x) W_{t+1}(x)\, dx = \int \mathbb{E}_w \psi(f_t(x, u_t, w)) W_t^{y_t}(x)\, dx \,. \qquad (12)$$

The above expression shows the dependance of the sequence $\{W_t\}$ on the control $u_{[0,T]}$ and the observation sequence $y_{[0,T]}$. Let this define the function $F_t$ as

$$W_{t+1} = F_t(W_t, u_t, y_t)\,. \qquad (13)$$

Let $\mathcal{W}$ be the set of all possible such functions $W_t$.

Via a standard dynamic programming argument, we can check that the Bellman return function $U$ is obtained by the recurrence relation

$$\forall W \in \mathcal{W}, \quad U_T(W) \;=\; \int M(x) W(x)\, dx\,, \qquad (14)$$

$$\forall W \in \mathcal{W}, \quad U_t(W) \;=\; \inf_u \mathbb{E}_y U_{t+1}\left(F_t(W, u, y)\right)\,. \qquad (15)$$

Moreover, assume that the minimum in $u$ is attained in (15) above at $u = \hat{\mu}_t(W)$. Then (13) and

$$u_t = \hat{\mu}_t(W_t) \qquad (16)$$

define an optimal controller for the stochastic control problem. The optimal cost is $U_0(N)$.

### 3.2.2 Minimax dynamic programming

Let us consider now the problem of minimizing $G(\mu)$. We have to introduce the *conditional state cost measure* and its cost density $W$ (according to the concepts introduced in section 2.1 following [1]). It is defined as the maximum possible past cost knowing the past information, as a function of current state. To be more precise, let us introduce the following subsets of $\Omega$. Given a pair $(u^t, y^t) \in \mathsf{U}^t \times \mathsf{Y}^t$, and a subset $A$ of $\mathbb{R}^n$, let

$$\Omega_t(A \mid u^t, y^t) = \{\omega \in \Omega \mid y^t = \eta^t(u^{t-1}, \omega^t), \text{ and } \phi_{t+1}(u^t, \omega^t) \in A\}\,. \qquad (17)$$

For any $x \in \mathbb{R}^n$, we shall write $\Omega_t(x \mid u^t, y^t)$, or simply $\Omega_t(x)$ when no ambiguity results, for $\Omega_t(\{x\} \mid u^t, y^t)$. And likewise for $\Omega_{t-1}(x)$.

The conditional cost measure of $A$ is $\sup_{\omega\in\Omega_{t-1}(A)}[N(x_0)+\Gamma(w_{[0,T]})]$, and hence the conditional cost density function is

$$W_t(x) = \sup_{\omega\in\Omega_{t-1}(x)} \left[\sum_{t=0}^{T-1} \Gamma_t(w_t) + N(x_0)\right] .$$

Initialize this sequence with

$$W_0(x) = N(x) .$$

It is a simple matter to write recursive equations of the form

$$W_{t+1} = F_t(W_t, u_t, y_t) .$$

In fact, $F_t$ is defined by the following. Let for ease of notations

$$\mathsf{Z}_t(x \mid u, y) = \{(\xi, v) \in \mathbb{R}^n \times \mathsf{W} \mid f_t(\xi, u, v) = x, \quad h_t(\xi, u, v) = y\} ,$$

then we have

$$W_{t+1}(x) = \sup_{(\xi,v)\in\mathsf{Z}_t(x|u_t,y_t)} [W_t(\xi) + \Gamma_t(v)] . \tag{18}$$

It is worthwhile to notice that, for any function $\psi(x)$, (such that the max exists)

$$\max_x [W_{t+1}(x) + \psi(x)] = \mathbb{F}_{w_t} \max_{x|h_t(x,w_t)=y} [W_t(x) + \psi(f_t(x, u_t, w_t))]$$

and that hence

$$\max_y \max_x [W_{t+1}(x) + \psi(x)] = \max_x \mathbb{F}_{w_t}[W_t(x) + \psi(f_t(x, u_t, w_t)] ,$$

the counterparts of (12) and (11) above.

As was probably first shown in [20], (also presented in a talk in Santa Barbara in july 1993), one can do simple dynamic programming in terms of this function $W$. The value function $U$ will now be obtained through the following relation

$$\forall W \in \mathcal{W}, \quad U_T(W) = \sup_x (M(x) + W(x)) , \tag{19}$$

$$\forall W \in \mathcal{W}, \quad U_t(W) = \inf_u \sup_y U_{t+1}(F_t(W, u, y)) . \tag{20}$$

Moreover, assume that the minimum in $u$ is attained in (20) above at $u = \hat{\mu}(W)$. Then it defines an optimal feedback (16), with $W_t$ now defined by (18), for the minimax control problem. The optimal cost is $U_0(N)$.

Of course, all our set up has been arranged so as to stress the parallel between (14),(15) on the one hand, and (19),(20) on the other hand.

## 3.3 Separation theorem

### 3.3.1 Stochastic separation theorem

We are here in the stochastic setup. The performance criterion is $H$ and $W$ stands for the conditional state probability density.

We introduce the *full information* Bellman return function $V_t$ defined by the classical dynamic programming recursion :

$$\forall x \in \mathbb{R}^n, \quad V_T(x) = M(x)\,,$$

$$\forall x \in \mathbb{R}^n, \quad V_t(x) = \inf_u \mathbb{E}_{w_t} V_{t+1}(f_t(x,u,w_t))\,.$$

Then we can state the following result.

**Proposition 1** *Let*

$$S_t(x,u) := \mathbb{E}_{w_t} V_{t+1}\left(f_t(x,u,w_t)\right) W_t(x)\,.$$

*If there exists a (decreasing) sequence of (positive) numbers $R_t$ with $R_T = 0$ such that,*

$\forall t \in [0, T-1], \forall u_{[0,T]} \in \mathbf{U}, \forall \omega \in \Omega,$

$$\int \min_u S_t(x,u)\,dx + R_t = \min_u \int S_t(x,u)\,dx + R_{t+1}\,,$$

*then the optimal control is obtained by minimizing the conditional expectation of the full information Bellman return function, i.e. choosing a minimizing $u$ in the right hand side above.*

**Proof** The proof relies on the following fact :

**Lemma 1** *Under the hypothesis of the proposition, the function*

$$U_t(W) = \int V_t(x) W(x)\,dx + R_t \tag{21}$$

*satisfies the dynamic programming equations (14)(15).*

Let us check the lemma. Assume that

$$\forall W_{t+1} \in \mathcal{W}, \quad U_{t+1}(W_{t+1}) = \int V_{t+1}(x) W_{t+1}(x)\,dx + R_{t+1}$$

and apply (15), using (12)

$$U_t(W_t) = \min_u \mathbb{E}_y \int \mathbb{E}_{w_t} V_{t+1}(f_t(x,u,w_t)) W_t^y(x)\,dx + R_{t+1}$$

and, according to (11) this yields

$$U_t(W_t) = \min_u \int \mathbb{E}_{w_t} V_{t+1}(f_t(x,u,w_t)) W_t(x)\, dx + R_{t+1}\,.$$

Using the hypothesis of the proposition and Bellman's equation for $V_t$, it comes

$$U_t(W_t) = \int V_t(x) W_t(x)\, dx + R_t\,,$$

and the recursion relation holds.

The hypothesis of the theorem sounds in a large extent like wishfull thinking. It holds, as easily checked, in the linear quadratic case. (In that case, symmetry properties result in the certainty equivalence theorem.) There is little hope to find other instances. We state it here to stress the parallel with the minimax case.

### 3.3.2 Minimax separation theorem

This section is based upon [6] [7]. The same result is to appear independantly in [17].

We are now in the minimax setup. The performance criterion is $G$, and $W$ stands for the conditional state cost density.

We introduce the *full information* Isaacs Value function $V_t(x)$ which satisfies the classical Isaacs equation:

$$\forall x \in \mathbb{R}^n, \quad V_T(x) = M(x)\,,$$

$$\forall x \in \mathbb{R}^n, \quad V_t(x) = \inf_u \mathbb{F}_{w_t} V_{t+1}(f_t(x,u,w_t))\,.$$

Notice that we do not need that the Isaacs condition, i.e. the existence of a saddle point in the right hand side above, hold. If it does not, $V$ is an upper value, which is what is needed in the context of minimax control.

It is convenient here to introduce a binary operation denoted $\oplus$ which can be either the ordinary addition or its dual in our morphism: the max operation.

**Proposition 2** *Let*

$$S_t(x,u) = \mathbb{F}_{w_t}[V_{t+1}\,(f_t(x,u,w)) + W_t(x)]\,.$$

*If there exists a (decreasing) sequence of numbers $R_t$, such that,*

$$\forall t \in [0, T-1], \forall u_{[0,T]} \in \mathbf{U}, \forall \omega \in \Omega,$$
$$\max_x \min_u S_t(x,u) \oplus R_t = \min_u \max_x S_t(x,u) \oplus R_{t+1}\,,$$

*then the optimal control is obtained by minimizing the conditional worst cost, future cost being measured according to the full information Isaacs Value function, i.e. taking a minimizing $u$ in the right hand side above.*

**Proof** The proof relies on the following fact :

**Lemma 2** *Under the hypothesis of the proposition, the function*

$$U_t(W) = \max_x [V_t(x) + W(x)] \oplus R_t$$

*satisfies the dynamic programming equations (19)(20).*

Let us check the lemma. Assume that

$$\forall W_{t+1} \in \mathcal{W}\,, \quad U_{t+1}(W_{t+1}) = \max_x [V_{t+1}(x) + W_{t+1}(x)] \oplus R_{t+1}$$

and apply (20), using (18)

$$U_t(W) = \min_u \max_y \left( \max_x [V_{t+1}(x) + \max_{(\xi,v)\in \mathsf{Z}_t(x|u,y)} (W_t(\xi) + \Gamma_t(v))] \oplus R_{t+1} \right) .$$

The max operations may be merged into

$$U_t(W) = \min_u \left( \max_{\xi,v} [V_{t+1}(f_t(\xi,u,v)) + \Gamma_t(v) + W_t(\xi)] \oplus R_{t+1} \right) .$$

Then, using the hypothesis of the proposition and Isaacs equation for $V$, it comes

$$U_t(W) = \max_x [V_t(x) + W_t(x)] \oplus R_t\,,$$

thus establishing the recursion relation.

The hypothesis of the proposition is not as unrealistic as in the stochastic case. It is satisfied in the linear quadratic case, but more generally, it can be satisfied if $S$ is convex-concave, for instance, with $\oplus$ the ordinary addition and $R_t = 0$ (or $\oplus$ the max operation and $R_t = -\infty$). Moreover, in that case, the same $u$ provides the minimum in both sides, yielding a certainty equivalence theorem.

## 3.4 An abstract formulation

It is known that in the stochastic control problem, some results, including derivation of the separation theorem, are more easily obtained using a more abstract formulation of the observation process, in terms of a family of $\sigma$-fields $\mathcal{Y}_t$ generated in the disturbance space. The axioms are that

- the brownian motion $w_t$ is *adapted* to the family $\mathcal{Y}_t$,
- the family $\mathcal{Y}_t$ is *increasing*.

The same approach can be pursued in the minimax case. Instead of an explicit observation through an output (2), one may define the observation process in the following way. To each pair $(u_{[0,T]}, \omega)$ the observation process associates a sequence $\{\Omega_t\}_{t\in[0,T]}$ of subsets of $\Omega$. The axioms are that, for any $(u_{[0,T]}, \omega)$, the corresponding family $\Omega_t$ satisfies the following properties.

- The process is *consistant*, i.e. $\forall t, \quad \omega \in \Omega_t$.
- The process is *strictly non anticipative*, i.e. $\omega \in \Omega_t \Leftrightarrow \omega^{t-1} \in \Omega_t^{t-1}$ where $\Omega_t^{t-1}$ stands for the set of restrictions to $[0, t-1]$ of the elements of $\Omega_t$.
- The process is *with complete recall*: $\forall (u_{[0,T]}, \omega), \quad t < t' \Rightarrow \Omega_t \supset \Omega_{t'}$.

In the case considered above, we have

$$\Omega_t = \Omega(\mathbb{R}^n \mid u^t, y^t)$$

but the abstract formulation suffices, and allows one, for instance, to extend the minimax certainty equivalence principle to a variable end time problem. See [7] for a detailed derivation.

One may think of the subsets $\Omega_t$ as playing the role of the measurable sets of the $\sigma$-field $\mathcal{Y}_t$.

# 4 The continuous time control problem

## 4.1 The problem

We now have a continuous time system, of the form

$$\dot{x} = f_t(x, u, w), \tag{22}$$

$$y = h(x, w). \tag{23}$$

The notations will be the counterpart of the discrete ones. In particular, $u^t$, will stand for the restriction to $[0, t]$ of the continuous time function $u_{[0,T]} : t \mapsto u_t$. We shall again let $\mathsf{U}^t$ designate the set of such segments of function. Likewise for $w^t \in \mathsf{W}^t$, $\omega^t \in \Omega^t$, and $y^t \in Y^t$. Notice however that (3) and (4) must be replaced by

$$x_t = \phi_t(u^t, \omega^t), \tag{24}$$

$$y_t = \eta_t(u^t, \omega^t). \tag{25}$$

and similarily for (5) and (6).

Admissible controllers will be of the form $u_t = \mu_t(u^t, y^t)$. This seems to be an implicit definition, since $u_t$ is contained in $u^t$. In fact, it is hardly more so than any feedback control. In any extent, we let $\mathcal{M}$ be the class of controllers of that form, such that they generate a unique trajectory for any $\omega \in \Omega$.

As in the discrete case, we may always bring a classical integral plus terminal cost to the form (7). The two problems we want to investigate are again the minimization of $H(\mu)$ given by 8 with a stochastic model for $\omega$ and that of $G(\mu)$ given by 9 with a cost density for $\omega$ (or its sole set membership description if we take this cost density constant).

## 4.2 Hamilton Jacobi theory

### 4.2.1 Stochastic Hamilton Jacobi theory

In the continuous time case, the technicalities of diffusion processes and Ito calculus make the stochastic problem much more complex than its discrete counterpart, or, for that matter, than its continuous minimax counterpart. As far as we know, the classical litterature concentrates on simpler, technically tractable, particular cases of the system (22),(23). Typically, classical nonlinear stochastic control deals with the system

$$dx_t = b_t(x,u)\,dt + \sigma_t(x,u)\,dw_t\,, \tag{26}$$

$$dy_t = c_t(x)\,dt + dv_t\,. \tag{27}$$

where $v_t$ and $w_t$ are standard independent vector brownian motions, and the above equations are to be taken in the sense of stochastic integrals. We shall need the notation $\sigma\sigma' = a$ where the prime stands for transposed, i.e.

$$a_{ij} = \sum_k \sigma_{ik}\sigma_{jk}\,.$$

Under suitable regularity and growth assumptions, one may compute a conditional state probability distribution $W_t$ through the stochastic PDE (which can be derived, for instance, from Zakai's equation, see [12]), the dual form of which may be writen $W_0 = N$ and, for any function $\psi(\cdot) \in C^2(\mathbb{R}^n)$,

$$d\int \psi(\xi)W_t(\xi)\,d\xi = \left(\int (L_t(u)\psi)(\xi)W_t(\xi)\,d\xi\right)dt + \left(\int \psi(\xi)W_t(\xi)[c_t'(\xi) - \bar{c}_t']\,d\xi\right)(dy_t - \bar{c}_t\,dt)\,,$$

where

$$(L_t(u)\psi)(\xi) = \frac{\partial\psi}{\partial x}(\xi)b_t(\xi,u) + \frac{1}{2}\sum_{ij}\frac{\partial^2\psi}{\partial x_i\partial x_j}(\xi)a_{ij}(\xi,u)\,, \tag{28}$$

and $\bar{c}_t$ stands for the conditional expectation of $c_t(x_t)$:

$$\bar{c}_t = \int c_t(z)W_t(z)\,dz\,.$$

A full information control problem can be written in terms of that probability density as a state. We refer to [12] for a complete treatment. The formal development is too intimately intermingled with the technical aspects to lend itself to a simple exposition of the kind given here. In particular, a nonlinear Hamilton Jacobi theory would imply Ito calculus with an infinite dimensional state, which we have rather avoid to write.

#### 4.2.2 Minimax Hamilton Jacobi theory

The minimax problem is not as complex as the stochastic one, at least to state formally, and as long as one only seeks sufficient conditions. It was independently developed in [6], and in [19] in a slightly less general context, but with a much more complete development in that it includes a first mathematical analysis of the resultant Isaacs equation.

We introduce the counterpart of (17) : for a given pair $(u^t, y^t) \in \mathsf{U}^t \times \mathsf{Y}^t$ and a subset $A$ of $\mathbb{R}^n$, let

$$\Omega_t(A \mid u^t, y^t) = \{\omega \in \Omega \mid y^t = \eta^t(u^t, \omega^t), \text{ and } \phi_t(u^t, \omega^t) \in A\}$$

be the conditional disturbance subset of $A$, and again write $\Omega_t(\xi)$ instead of $\Omega_t(\{\xi\} \mid u^t, y^t)$. The conditional cost density function is now

$$W_t(x) = \sup_{\omega \in \Omega_t(x)} \left( N(x_0) + \int_0^T \Gamma_t(w_t)\, dt \right) .$$

If it is $C^1$, $W_t$ satisfies a forward hamilton Jacobi equation. Let

$$\mathsf{W}_t(x \mid y) = \{w \in \mathsf{W} \mid h_t(x, w) = y\} ,$$

then this forward equation is, for $u^t$ and $y^t$ fixed:

$$\frac{\partial W_t(x)}{\partial t} = \sup_{w \in \mathsf{W}_t(x|y_t)} \left[ -\frac{\partial W(x)}{\partial x} f_t(x, u_t, w) + \Gamma_t(w) \right] \tag{29}$$

which we write as

$$\frac{\partial}{\partial t} W_t = \mathbb{F}_w^y \left[ -\frac{\partial W(x)}{\partial x} f_t(x, u_t, w) \right] =: F_t(W_t, u_t, y_t)$$

and, together with the initial condition $W_0 = N$, it may define $W_t$ along any trajectory.

Assume $\mathcal{W}$ is endowed with a topology for which $U$ is absolutely continuous in $W$, and admits a Gâteaux derivative $D_W U$. Then, the value function $U_t(W)$ is obtained through the following Isaacs equation. $U_T$ is again given by (19), and

$$\forall W \in \mathcal{W}, \quad \frac{\partial U_t(W)}{\partial t} + \inf_{u \in \mathsf{U}} \sup_{y \in \mathsf{Y}} D_W U_t(W) F_t(W, u, y) = 0 \,. \tag{30}$$

Moreover, assume that the minimum in $u$ is attained in (30) above at $u = \hat{\mu}_t(W)$, then (16) defines an optimal feedback for the minimax control problem. The optimal cost is $U_0(N)$.

Notice again that the easy task is to show a *sufficient* condition: if there exist $C^1$ functions $W$ and $U$ satisfying these equations, and if the feedback (16) is admissible, then we have a solution of the problem. It is worth noticing that the only existence result we are aware of is in [18], and is in a particular case somewhat similar to the set up we have outlined for the stochastic case.

A further remark is that, in a case, say, where $N = 0$ and $\Gamma = 0$, the function $W$ only, and exactly, characterizes the reachable set given the past information. Let $\mathsf{X}_t(u^t, y^t)$ be that set, then we have

$$W_t(x) = \begin{cases} 0 & \text{if } x \in \mathsf{X}_t\,, \\ -\infty & \text{if } x \notin \mathsf{X}_t\,. \end{cases}$$

This is of course highly nondifferentiable. An apparent serious drawback for this theory, since this is an important case.

There are two ways that may help resolve this problem. The first one is developed in [6]. It consists in using the Fenchel transform $W^*$ of $W$, defined as

$$W^*(p) = \min_x [(p, x) - W(x)]\,. \tag{31}$$

We show that, under some additional assumptions, $W^*$ satisfies a dual forward Hamilton Jacobi equation:

$$\frac{\partial W_t^*(p)}{\partial t} + \sup_{w \in \mathsf{W}_t(\xi_t | y_t)} [-p f_t(\xi_t, u_t, w) + \Gamma_t(w)] = 0\,.$$

where

$$\xi_t = \frac{\partial W_t^*(p)}{\partial p}\,.$$

Now, $U$ can be taken as a function of $W^*$. If $W$ is a concave function, i.e. if $\mathsf{X}_t$ is convex in the case (31) above, the dual approach yields the exact minimax control. If $W$ is not concave, the strategy thus computed yields a guaranteed cost $U_0(N^*)$.

Another possible way around the non differentiability of $W$ is given by the following remark. One can replace $W$ in the theory by a *parametrization* of $W$. Let $\mathcal{P}$ be a topological space, called the parameter space, $\pi : \mathcal{P} \to \mathcal{W}$ be a one to one map. Assume that to any pair $(u^t, y^t)$ we can asociate a time function $p_t$ satisfying a differential equation

$$\dot{p}_t = \mathcal{F}_t(p_t, u_t, y_t)$$

such that $\pi(p_t)$ be the conditional cost density $W_t$ of the process. Then it is clear that the Value function can be expressed in terms of $p$ instead of $U$, and we recover the necessary differentiability to write the equivalent of (30), which becomes

$$\forall p \in \mathcal{P}, \quad \frac{\partial U_t(p)}{\partial t} + \inf_{u \in \mathsf{U}} \sup_{y \in \mathsf{Y}} D_p U_t(p) \mathcal{F}_t(p,u,y) = 0\,.$$

In the case (31), $p$ parametrizes as well $\mathsf{X}_t(u^t, y^t)$ as its characteristic fuction $W$. This is what we do in [22], where it is clear that $W$ (or $\mathsf{X}_t$) lies on a three dimensional manifold of $\mathcal{W}$ (or $2^{\mathsf{X}}$), so that we may take $\mathbb{R}^3$ for $\mathcal{P}$.

**The $L_t(u)$ operator** We notice here a strange parallel. Assume that the dynamics are as in (26), or more precisely, since $w$ is not a white noise anymore but a (deterministic) decision variable

$$f_t(x,u,w) = b_t(x,u) + \sigma_t(x,u)w\,.$$

Notice that the natural dual to the Gaussian law is the cost measure $\Gamma(w) = -1/2\|w\|^2$. Then, a counterpart of (28) is, for a function $\psi(x(t))$

$$\mathbb{F}\left[\frac{d\psi}{dt}(\xi,u)\right] = \frac{\partial \psi}{\partial x}(\xi) b_t(\xi,u) + \sum_{i,j} \frac{\partial \psi}{\partial x_i}(\xi) \frac{\partial \psi}{\partial x_j}(\xi) a_{ij}(\xi,u)\,.$$

At this stage, we do not know wheher the similarity with (28) is anything else than a curiosity. Notice that the above operator is not linear. (Here we refer to (max, +) linearity.)

## 4.3 Separation theorem

### 4.3.1 Stochastic separation theorem

We may take advantage of the linear character of the equation (21) to write, at least formally, the continuous time counterpart to the stochastic separation principle of section 3.3.1. We need first introduce the *full information* (state feedback) Bellman function $V_t(x)$ which satisfies the *stochastic Bellman equation* (see [13])

$$V_T = M\,,$$

$$\forall (t,x) \in [0,T] \times \mathbb{R}^n, \quad \frac{\partial V_t}{\partial t}(x) + \inf_u (L_t(u)V_t)(x) = 0\,.$$

We can then state the following result.

**Proposition 3** *Let*

$$S_t(x,u) = (L_t(u)V_t)(x) W_t(x)\,.$$

*If there exists a real (positive) $L^1([0,T])$ function $r_t$ such that,*

$\forall t \in [0,T], \forall u_{[0,T]} \in \mathbf{U}$, *almost surely*

$$\int \min_u S_t(x,u)\,dx + r_t = \min_u \int S_t(x,u)\,dx\,,$$

*then an optimal control is obtained by choosing the minimizing $u$, that we shall call $\hat{\mu}_t(W_t)$, in the right hand side above.*

**Proof**. The proof relies on the following fact

**Lemma 3** *Under the hypothesis of the proposition, the stochastic process $\alpha_t = U_t(W_t)$ is a submartingale for any admissible control, and a martingale if $u_t = \hat{\mu}_t(W_t)$, where the function $U_t$ is defined over the set $L^1(I\!R^n)$ by*

$$U_t(W) = \int_{I\!R^n} V_t(\xi)W(\xi)\,d\xi + R_t\,,$$

*and*

$$R_t = \int_t^T r_s\,ds\,. \tag{32}$$

Let us check the lemma. Consider the diffusion process $\alpha_t = U_t(W_t)$ where the system, and thus the filter, is driven by a control process $u_t$. It satisfies the stochastic differential equation

$$d\alpha_t = \int [L_t(u_t)V_t - \inf_u L_t(u)V_t](\xi)W_t(\xi)\,d\xi\,dt - r_t\,dt \\ + \int V_t(\xi)W_t(\xi)[c'_t(\xi) - \bar{c}'_t]\,d\xi[dy_t - \bar{c}_t\,dt]\,.$$

Assume that $u_t$ is *admissible*, i.e. measurable over the $\sigma$-field $\mathcal{Y}_t$ generated by the observation process $y_t$, and take the conditional expectation. One obtains, at least formally

$$I\!E^{\mathcal{Y}_t} d\alpha_t = \int [L_t(u_t)V_t - \inf_u L_t(u)V_t](\xi)W_t(\xi)\,d\xi\,dt - r_t\,dt\,.$$

It follows that if $u_t = \hat{\mu}_t(W_t)$, which is indeed admissible, the hypothesis of the proposition yields $I\!E d\alpha_t = 0$, and for any other admissible control $I\!E d\alpha_t \geq 0$.

Thus under the feedback control $\hat{\mu}_t(W_t)$, $I\!E U_T(W_T) = U_0(W_0)$, hence, recalling that $V_T = M$ and $W_0 = N$, $I\!E M(x_T) = I\!E V_0(x_0) + R_0$. And for any other admissible control, $I\!E U_T(W_T) \geq U_0(W_0)$, hence $I\!E M(x_T) \geq I\!E V_0(x_0) + R_0$.

The above proof is formal in that we have not detailed the regularity and growth hypotheses under which these calculations are valid. But it can be made rigorous, and provide a proof of the separation theorem for the linear quadratic case for instance.

### 4.3.2 Minimax separation theorem

Introduce as in the discrete time case the *full information* Isaacs'Value function $V$. It satisfies the Isaacs equation

$$\begin{aligned} \forall x \in \mathbb{R}^n, \quad V_T(x) &= M(x)\,, \\ \forall t, \forall x \in \mathbb{R}^n, \quad \frac{\partial V_t(x)}{\partial t} &= \min_{u \in \mathsf{U}} \mathbb{F}_{w_t} \left( \frac{\partial V_t(x)}{\partial x} f_t(x,u,w_t) \right) . \end{aligned}$$

The use of weak solutions, the viscosity solution, is now well understood. However, for our purpouse here, which is to stress the formal duality according to Quadrat's morphism, we shall assume that $V$ and $W$ are $C^1$. We shall also assume that the full information game admits a unique state feedback solution $u_t = \phi_t^*(x_t)$, argument of the min above.

As in [3], introduce also the *auxiliary problem*

$$\max_{x \in \mathbb{R}^n} [V_t(x) + W_t(x)]\,.$$

and assuming it has a (nonunique) solution, let $\hat{X}_t$ be the set of maximizing $x$'s, or *conditional worst states*.

We have the following fact:

**Proposition 4** *Let*

$$S_t(x,u) = \mathbb{F}_{w_t} \left[ \frac{\partial V_t}{\partial x}(x) f_t(x,u,w_t) \right] .$$

*If there exists a real (positive) $L^1([0,T])$ function $r_t$ such that*

$$\forall t \in [0,T], \forall u_{[0,T]} \in \mathbf{U}, \forall \omega \in \Omega, \quad \min_{x \in \hat{X}_t} \min_u S_t(x,u) + r_t = \min_u \max_{x \in \hat{X}_t} S_t(x,u)\,,$$

*then an optimal control is obtained by minimizing the conditional worst rate of increase of the full information Value function among the conditional worst states, i.e. taking the minimizing $u$ in the right hand side above.*

**Proof** The proof relies on the following fact :

**Lemma 4** *Under the hypothesis of the proposition, the function*

$$U_t(W) = \max_x [V_t(x) + W_t(x)] + R_t$$

*with $R_t$ defined as in (32) satisfies the dynamic programming equations (19), and (30) replacing derivatives with right derivatives in time.*

The lemma hinges on Danskin's theorem [11] to get for the right time derivative

$$\left(\frac{\partial U_t}{\partial t}\right)^+ (W) = \max_{x\in\hat{X}_t} \frac{\partial V_t}{\partial t}(x) - r_t$$

and for the directional derivative in a direction $dW \in \mathcal{W}$:

$$D_W U_t(W) \cdot dW = \max_{x\in\hat{X}_t} dW(x)\,.$$

Then it is a simple matter to place this in (30), notice that because all $x \in \hat{X}_t$ maximize the auxiliary problem, then at these points $-\partial W_t/\partial x = \partial V_t/\partial x$, and that, as in the case of mathematical expectations, the cascade of the two max operators $\max_y \max_{w\in W(x|y)}$ collapses in $\max_w$, to get the result.

**Remarks** The condition of the proposition looks a bit odd. A first remark is that $W_t$, hence $\omega$, *seems* not to enter it. This is of course *not* the case, because $\hat{X}_t$ depends on $W_t$.

The second remark is that we have quoted the proposition that way to stress a parallel with the other cases. (The parallel would have been better if we had not converted $-\max(-\bullet)$ in $\min(\bullet)$.) But its only reasonable use seems to be the following corollary, the now well known minimax certainty equivalence principle of [3],[9] :

**Corollary 1** *If* $\forall t \in [0,T]$, $\forall u_{[0,T]} \in \mathbf{U}$, $\forall \omega \in \Omega$, $\hat{X}_t$ *is a singleton* $\{\hat{x}_t\}$*, then an optimal control is obtained by replacing* $x_t$ *by* $\hat{x}_t$ *in the optimal state feedback of the full information problem, i.e. taking* $u_t = \phi_t^*(\hat{x}_t)$*.*

## 4.4 An abstract formulation

The abstract formulations of the obsrvation process have indeed be originally introduced for the continuous time problems. The parallel here is exactly the same as in the discrete time case, the only difference for the minimax problem being that nonanticipativeness of the process is now written as

$$\omega \in \Omega_t \Leftrightarrow \omega^t \in \Omega_t^t\,.$$

This approach to proving the certainty equivalence theorem was first proposed in [9]. It allows one to extend the theorem to variable end time problems.

# 5 Conclusion

The parallel between stochastic and minimax control appears thus as striking, even if some technicalities make it less clear in the continuous time case than in the discrete time case. Some more work probably remains to be done to fully explain and exploit it. But it is clear that "Quadrat's morphism" is at the root of the problem.

# REFERENCES


[1] M. Akian, J-P. Quadrat and M. Viot: "Bellman Processes" 11th International Conference on Systems Analysis and Optimization, Sophia Antipolis, France, 1994.

[2] F. Baccelli, G. Cohen, J-P. Quadrat, G-J. Olsder: *Synchronization and Linearity*, Wiley, Chichester, U.K., 1992.

[3] T. Başar and P. Bernhard: *$H_\infty$ Optimal Control and Related Minimax Design Problems, a Game Theory Approach*, Birkhauser, Boston, Mas 1991.

[4] J.A. Ball and J.W. Helton : "$\mathcal{H}_\infty$-Control for Nonlinear Plants: Connections with Differential Games", 28th CDC, Tampa, Fla, 1989

[5] J.A. Ball, J.W. Helton and M.L. Walker, "$H_\infty$-Control For Nonlinear Systems with Output Feedback", IEEE Trans. Automatic Control, **AC-38**, 1993, pp 546–559.

[6] P. Bernhard "Sketch of a Theory of Nonlinear Partial Information Min-Max Control", Research Report 2020, INRIA, France, 1993, to appear in a revised form, IEEE Trans. A.C., 1994.

[7] P. Bernhard: "A Discrete Time Min-max Certainty Equivalence Principle", to appear, Systems and Control Letters, 1994.

[8] P. Bernhard and G. Bellec : "On the Evaluation of Worst Case Design, with an Application to the Quadratic Synthesis Technique", 3rd IFAC symposium on Sensitivity, Adaptivity and Optimality, Ischia, Italy, 1973

[9] P. Bernhard and A. Rapaport: "Min-max Certainty Equivalence Principle and Differential Games", revised version of a paper presented at the Workshop on Robust Controller Design and Differential Games, UCSB, Santa Barbara, California, 1993. Submitted for publication.

[10] G. Didinsky, T. Başar, and P. Bernhard "Structural Properties of Minimax Policies for a Class of Differential Games Arising in Nonlinear $H_\infty$-control", Systems and Control Letters, 1993, pp

[11] F. Clarke: *Optimization and Nonsmooth Analysis*, John Wiley, 1983.

[12] W.H. Fleming and E. Pardoux: "Existence of Optimal Controls for Partially Observed Diffusions", SIAM Jal on Control and Optimisation **20**, 1982, pp 261–283.

[13] W.H. Fleming and R.W. Rishel: *Deterministic and Stochastic Optimal Control*, Springer Verlag, New York, 1975

[14] V. Fromion "A Necesary and Sufficient Condition for Incremental Stability of Nonlinear Systems", to appear in this symposium.

[15] A. Isidori : "Feedbak Control of Nonlinear Systems" 1st ECC, Grenoble, France, 1991

[16] A. Isidori and A. Astolfi : Disturbance Attenuation and $\mathcal{H}_\infty$ Control via Measurement Feedback in Nonlinear Systems", IEEE trans. Automatic Control, **AC 37**, pp1283-1293, 1992

[17] M. James "On the Certainty Equivalence Principle and the Optimal Control of Partially Observed Dynamic Games", to appear IEEE Trans. on Automatic Control, 1994.

[18] M.R. James and J.S. Baras "Partially Observed Differential Games, Infinite Dimensional HJI Equations, and Nonlinear $H_\infty$ Control". To appear.

[19] M.R. James, J.S. Barras and R.J. Elliott: "Output Feedback Risk Sensitive Control and Differential Games for Continuous Time Nonlinear Systems", 32nd CDC, San Antonio, Texas, 1993

[20] M.R. James, J.S. Barras and R.J. Elliott: "Risk Sensitive Control and Dynamic Games for Partially Observed Discrete Time Nonlinear Systems", IEEE Trans. on Automatic Control **AC-39**, 1994, pp 780–792

[21] J-P. Quadrat: "Théorèmes asymptotiques en programmation dynamique", *Compte Rendus de l'Académie des Sciences*, 311: pp 745–748, 1990.

[22] A. Rapaport and P. Bernhard: "Un jeu de poursuite-evasion avec connaissance imparfaite des coordonnées", International Symposium on Differential Games and Applications, St Jovite, Canada, 1994.

[23] A.J. van der Schaft : "Nonlinear State Space $\mathcal{H}_\infty$ Control Theory", ECC 1993, Groningen, Nederlands