Discrete and Continuous Time Partial Information Minimax Control

Pierre Bernhard

July 1994

Abstract

We review the current state of partial information minimax control, stressing the parallel with stochastic control.

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_∞-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, but many others have followed suit : [13],[14],[21], [4], [5] and more recently [16],[17]. 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_∞-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_∞-optimal control. We dont stress quadratic loss functions. But more importantly, we claim that the minimax problem is only an intermediary step in H_∞ 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, ΔP the model uncertainty). In that respect, the nonlinear equivalent is *not* the minimax problem usually considered, but rather the contraction problem independently tackeled by [12].

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 [19] [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 [19] [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. Let us briefly review some concepts, based on [1].

2.1 Cost measure

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

- $K(\Omega) = 0$
- $K(\emptyset) = -\infty$
- 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).

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

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

One has the following theorem (Akian)

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

2.2 Cramer transform

The parallel between the ordinary and the $(\max, +)$ algebra is often based upon large deviation theory in probabilities. It is then based upon the fact that the function

$$\lim_{a \to \infty} \log_a(a^x)$$

turns sum into max and product into sum by the identities

$$\lim_{a \to \infty} \log_a(a^x + a^y) = \max(x, y)\,, \quad \log_a(a^x a^y) = x + y\,.$$

The above morphism is fully exploited by the *Cramer transform* $\mathcal{C}$ from the set of positive measures to that of proper, l.s.c. convex functions defined in terms of the classical Fenchel transform $\mathcal{F}$ and Laplace transform $\mathcal{L}$ by $\mathcal{C} = \mathcal{F} \circ \ln \circ \mathcal{L}$. Since we insist on using the $(\max, +)$ algebra instead of the $(\min, +)$ algebra, we shall rather make reference to the *opposite* of the Cramer transform.

Let f and g be transformed by the opposite Cramer transform into ϕ and ψ, then, their convolution $f \star g$ is taken into the sup-convolution $\phi \nabla \psi(p) = \sup_q(\phi(q) + \psi(p - q))$ of ϕ and ψ. Let also $\bar{g}(t) = g(-t)$, it is transformed into $\bar{\psi}(x) = \psi(-x)$. Thus $f * \bar{g}$ is transformed into $\phi \nabla \bar{\psi}$.

This last remark introduces a (purely formal) parallel between the L^2 scalar product

$$\int f(x) g(x)\, dx = (f * \bar{g})(0)$$

and what is called the $(\max, +)$ scalar product

$$\sup_x (\phi(x) + \psi(x)) = (\phi \nabla \bar{\psi})(0)$$

which is, for that reason, often denoted as a scalar product (ϕ, ψ) by Baras and James and others.

The opposite Cramer transform also turns the product *by a constant* k into a sum as $-\mathcal{C}(kf) = \ln k + \phi$ (unfortunately this does not carry over to product of functions), and introduces a direct relationship between $m = \int f(x)\, dx$ and $\sup_x \phi(x) = \ln(m)$. Hence, if f is the density of a probability measure, $m = 1$, and thus $\sup_x \phi(x) = 0$ making ϕ the density of a cost measure.

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

$$x_t = \phi_t(u_{[0,T]}, \omega)$$
$$y_t = \eta_t(u_{[0,T]}, \omega)$$

Finally, we shall call u^t a partial sequence $(u_0, u_1, \ldots, u_t)$ and 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 of the form $u_t = \mu_t(u^{t-1}, y^{t-1})$, i.e. strictly causal. We denote by $\mathcal{M}$ the class of such controllers.

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

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

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

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 disturbance".

We want to compare here two ways of turning this unprecise statement into a meaningful mathematical problem. In the first one, *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. The criterion to be minimized over $\mathcal{M}$ is then

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

In the second case, we do not modelize the perturbation otherwise than through the data of the set Ω, and we wish to choose μ in such a way as to get the best possible guaranteed payoff, i.e. minimizing

$$G(\mu) = \sup_{\omega \in \Omega} J(\mu, \omega)\,.$$

Remark 1 *We get a better parallel if we replace J in G above by*

$$\bar{J}(\mu, \omega) = J(\mu, \omega) + N(x_0)\,. \tag{9}$$

and consequently, G by

$$\bar{G}(\mu) = \sup_{\omega \in \Omega} \bar{J}(\mu, \omega)\,. \tag{10}$$

Of course, N can as well be absorbed in L_0, however, this formulation preserves the symmetry with the continuous time case, and is natural to display explicitly the "cost" associated to the choice of x_0 by Nature. In the wording of [1], this associates a initial state cost density exactly dual to the initial state probability density of stochastic control. Notice however that the ensuing performance index is no longer purely terminal.

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

$$W_0(x) = N(x) \tag{11}$$

at each step, W_t^η 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 ψ,

$$\mathbb{E}_y \int \psi(x) W_t^y(x)\,dx = \int \psi(x) W_t(x)\,dx\,. \tag{12}$$

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:

$$\int \psi(x) W_{t+1}(x)\,dx = \int \mathbb{E}_w \psi(f_t(x, u_t, w)) W_t^{y_t}(x)\,dx \tag{13}$$

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

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

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

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

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

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

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

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

defines an optimal feedback 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 $\bar{G}(\mu)$. 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 Ω. Given a pair $(u^t, y^t) \in \mathsf{U}^t \times \mathsf{Y}^t$, and a subset A of $\mathbb{R}^n$, let

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

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 $\sup_{\omega \in \Omega_{t-1}(A)} N(x_0)$, and hence the conditional cost density function is

$$W_t(x) = \sup_{\omega \in \Omega_{t-1}(x)} N(x_0)\,.$$

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

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

then we have

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

As was probably first shown in [18], (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

$$\begin{aligned} \forall W \in \mathcal{W}, \quad U_T(W) &= \sup_x (M(x) + W(x)) && (19) \\ \forall W \in \mathcal{W}, \quad U_t(W) &= \inf_u \sup_y U_{t+1}\left(F_t(W, u, y)\right) && (20) \end{aligned}$$

Moreover, assume that the minimum in u is attained in (20) above at $u = \hat{\mu}(W)$. Then it defines an optimal feedback (16) 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 :

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

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

Then we can state the following result.

Proposition 1 *Let*

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

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 made, we have*

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

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 (13)

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

and, according to (12) this yields

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

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 [15].

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:

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

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

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*

$$S_t(x,u) = \max_w \left[V_{t+1}\left(f_t(x,u,w)\right) + W_t(x)\right]$$

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 made, we have*

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

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)

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

The max operations may be merged into

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

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 σ-fields $\mathcal{F}_t$ generated in the disturbance space. The axioms are that this family is

- *adapted* to the underlying brownian motion w_t,
- *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 Ω. The axioms are :

- 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 Ω_t^{t-1} stands for the set of restrictions to $[0, t-1]$ of the elements of Ω_t.

- The process is *with complete recall*, i.e.

$$\forall (u_{[0,T]}, \omega), \quad t < t' \Rightarrow \Omega_t \supset \Omega_{t'} \,.$$

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 Ω_t as playing the role of the measurable sets of the σ-field $\mathcal{F}_t$.

4 The continuous time controm 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 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), or (9) for the minimax problem. The two problems we want to investigate are again the minimization of $H(\mu)$ with a stochastic model for ω and that of $\bar{G}(\mu)$ with the sole set description of ω.

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 filtering deals with the system

$$\begin{aligned} dx &= f_t(x,u)\,dt + g_t(x)\,dw\,, \\ dy &= h_t(x)\,dt + dv\,. \end{aligned}$$

Then, under some regularity conditions, an unnormalized version of the conditional state probability density can be computed through a stochastic PDE (Zakai's equation), and a Hamilton Jacobi theory may be written in terms of that probability density as a state. We refer to [11] 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.

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 [17] 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 $\Omega_t(\xi)$ for $\Omega_t(\{\xi\} \mid u^t, y^t)$. The conditional cost density function is now

$$W_t(x) = \sup_{\omega \in \Omega_t(\xi)} N(x_0)\,.$$

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:

$$\frac{\partial W_t(x)}{\partial t} = \sup_{w \in \mathsf{W}_t(x|y_t)} \left[-\frac{\partial W(x)}{\partial x} f_t(x, u_t, w) \right] \tag{26}$$

which we write as

$$\frac{\partial}{\partial t} W_t = F_t(W_t, u_t, y_t)$$

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

Then, the value function $U_t(W)$ is obtained through the following Isaacs equation. U_T is again given by (19), and

$$\forall W \in \mathcal{W}, \quad \frac{\partial U_t(W)}{\partial t} + \inf_{u \in \mathsf{U}} \sup_{y \in \mathsf{Y}} D_W U_t(W) F_t(W, u, y) = 0 \tag{27}$$

Moreover, assume that the minimum in u is attained in (27) 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 [16], 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 $N = 0$, 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

$$W^*(p) = \min_x [(p, x) - W(x)]\,. \tag{28}$$

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

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

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 X_t is convex in the case (28) 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 (27), 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 (28), p parametrizes as well $\mathsf{X}_t(u^t, y^t)$ as its characteristic fuction W. This is what we do in [20], where it is clear that W (or X_t) lies on a three dimensional manifold of $\mathcal{W}$ (or 2^{X}), so that we may take $\mathbb{R}^3$ for $\mathcal{P}$.

4.3 Separation theorem

4.3.1 Stochastic separation theorem

The continuous time equivalent of the discrete time stochastic separation theorem above can be stated in a comparable manner, made more involved by the Ito calculus —although the linear character of (21) helps in that matter. However, again the technical dificulties all but prevent one to state it in a concise way. And the conditions required look even stranger than in the discrete time case. We do not attempt this feat here.

4.3.2 Minimax separation theorem

As a matter of fact, this theory yields a certainty equivalence theorem rather than a separation principle, as we shall see. It was first reported in [3].

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

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

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*

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

We have the following fact:

Proposition 3 *If the auxiliary problem admits for all $(u_{[0,T]}, \omega)$ and for all t a unique solution $\hat{x}_t$ (depending, of course, on (u^t, y^t) through W_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 $u_t = \phi_t^*(\hat{x}_t)$.*

Proof The proof relies on the following lemma

Lemma 3 *Under the hypothesis of the proposition, we have*

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

Checking the lemma is easy. It hinges on Danskin's theorem, which insures that the above formula for U results in

$$\frac{\partial U_t(W)}{\partial t} = \frac{\partial V_t(\hat{x}_t)}{\partial t}$$

and also, for any $dW \in \mathcal{W}$,

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

Once this is recognised, it is a simple matter to place these formulas in (27) which becomes, using (26),

$$\frac{\partial V_t(\hat{x}_t)}{\partial t} + \inf_{u\in\mathsf{U}} \sup_{y\in\mathsf{Y}} \sup_{w\in\mathsf{W}_t(x|y)} \left[-\frac{\partial W_t(\hat{x}_t)}{\partial x} f_t(\hat{x}_t, u, w) \right] = 0\,.$$

The two suprema merge in a sup over $w \in \mathsf{W}$. And because $\hat{x}_t$ solves the auxiliary problem over $\mathrm{I\!R}^n$,

$$-\frac{\partial W_t(\hat{x}_t)}{\partial x} = \frac{\partial V_t(\hat{x}_t)}{\partial x}$$

so that the above equation is just Isaacs equation for V, and thus satisfied. The same remark yields the fact that the minimizing u in (27) is just $\phi^*(\hat{x}_t)$.

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.

4.5 Filter design

In the continuous time problem, we can stress a further parallel between the stochastic problem and the minimax one.

In the linear case, one of the most famous results of classical stochastic system theory is the fact that the conditional expected value $\hat{x}_t$ of the state can be computed in real time via the celebrated Kalman filter.

It was shown independantly in [21] and in [10] that in the case where the output disturbance is an additive, i.e., (23) is replaced by

$$y_t = h_t(x_t) + v_t$$

and if $\hat{x}_t$ above is well defined and the certainty equivalence theorem holds, one can give a filter to compute $\hat{x}$. The situation is much less favourable than in the linear Kalman filter case, because the gains cannot, at this time, be computed independantly of the general, infinite dimensional, theory. But yet, its form is worthwhile noticing.

Having absorbed the integral term in the criterion somewhat confuses the issues here. In particular it obliges us to reintroduce v in the dynamics. We therefore write them as

$$\dot{x}_t = f_t(x_t, u_t, v_t, w_t)\,.$$

Then, it can be shown that, upon playing the optimal, certainty equivalent control (16), we have

$$\begin{aligned}\dot{\hat{x}}_t = {} & f(\hat{x}_t, \hat{\mu}_t(W_t), \hat{v}_t, \hat{w}_t) + \\ & \left[\frac{\partial^2 V_t}{\partial x^2} + \frac{\partial^2 W_t}{\partial x^2}\right]^{-1} \left(\frac{dh_t}{dx}\right)^t \left(\frac{\partial f_t}{\partial v}\right)^t \left(\frac{\partial W_t}{\partial x}\right)^t [\hat{x}_t, \hat{\mu}_t(W_t), y - h_t(\hat{x}_t), \hat{w}_t]\end{aligned}$$

Here, the upper index t stands for "transposed", $\hat{v}_t$ and $\hat{w}_t$ stand for the worst state feedback disturbances (maximizing in the full information Isaacs equation), and the last square bracket in the r.h.s. stands for the arguments of the whole r.h.s. It is a simple matter to check that in the linear quadratic case, one recovers the classical H_∞ filter.

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 remains to be done probably 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 IFAC/INRIA 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_∞ *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_∞-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_∞-control", Systems and Control Letters, 1993, pp

[11] 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.

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

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

[14] 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

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

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

[17] 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

[18] 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

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

[20] 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.

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