A lecture on the game theoretic approach to H_∞ optimal control

Pierre Bernhard
INRIA-Sophia Antipolis,
France

August 21, 1991

1 Introduction : The problems considered

1.1 Control design problems with H_∞ bounds

The theory of so-called H_∞-optimal control has received much attention in the last decade. It started with Zames' minimum sensitivity and mixed sensitivity problems in the early 80's, [11]. The main idea is as follows.

Given a plant $y = Gu$ (in the classical problem, it is assumed linear stationary, but we shall not need stationarity), devise a feedback control $u = Ky$ which shall make the *sensitivity transfer function* $T = (I + GK)^{-1}$ small, in order to be robust, i.e. resistant against model errors in G. The natural meaning of "small", to engineers, refers to the maximum value of (the norm of) $T(i\omega)$ over all frequencies, i.e. the H_∞ norm of T, which has to lie in the Hardy space H_∞ in order for the closed loop system to be stable. Further, for measurement noise insensivity, it would be desirable that the *complementary sensitivity function* $T_c = GK(I+GK)^{-1}$ also be small. However, since $T + T_c = 1$, it is impossible to make both small at a time. The solution to this dilemma is to realize that modelling errors introduce *low frequency* disturbances, while measurement errors tend to be *high frequency*. Hence the idea of *loop shaping* whereby one attempts to control the magnitude of T, say, at low frequencies, and of T_c at high frequencies. This is usually done by defining weighting functions $W(s)$ and $W_c(s)$, and controlling the norm of the combined transfer function $H = [WT \quad W_cT_c]$.

Building on that idea, many more refined problems have been defined, attempting also, for instance, to control the input activity transfer function $K(I+GK)^{-1}$, and so on. Provided that the various weighting functions be taken as rational, sometimes with care to the properness of the various systems involved, all these problems may be cast into the *standard problem* described below.

We do not have space to discuss here in more detail the merits of these approaches, nor shall we try to give credit to the contributors to this theory since it began. We refer the reader to the references [8], [7] and [9]. We shall, however, offer in the concluding chapter a slightly different rationale for the current approach, still related to robustness, but closer to the min-max approach we adopt here.

As a matter of fact, the problem was first tackled by operator theoretic techniques. It was only in 1989, after it was shown in [7] (appeared at the CDC of december 1988) that the standard solution could be expressed in terms of familiar Riccati equations, that it was fully understood that the problem could be simply stated in terms of a min-max linear quadratic problem. This brought about renewed interest in the old "worst case design" approach, allowing us to give non stationary versions of the theory, and to solve many open problems.

The following lecture is based upon the book [2], where a more detailed discussion of this aspect of the theory can be found.

1.2 The standard problem

We consider a two input two output linear system of the form

$$\begin{aligned} z &= G_{11}w + G_{12}u\,, \\ y &= G_{21}w + G_{22}u\,, \end{aligned}$$

where $u(t) \in \mathbf{R}^m$, $w(t) \in \mathbf{R}^\ell$, $y(t) \in \mathbf{R}^p$, $z(t) \in \mathbf{R}^q$. The system may be either in continuous time : $t \in \mathbf{R}$ or discrete time : $t \in \mathbf{Z}$. In the classical problem, it is taken stationary, thus G_{ij} may stand for a (linear, causal, stationary) operator, or for its transfer function.

We want to design a (causal) controller $u = Ky$, taken linear and stationary in the classical problem. Under this control, we get a linear system from w to z : $z = T_K w$, with a transfer function

$$T_K = G_{11} + G_{12}(I - KG_{22})^{-1}KG_{21} = G_{11} + G_{12}K(I - G_{22}K)^{-1}G_{21}.$$

We require that K stabilizes the system. Thus T_K will be in H_∞, meaning that the norm $\|T_K(z)\|$ is bounded, in the right half plane in the continuous time case, outside the unit disk in the discrete time case. Moreover, T_K may be extended to the boundary of this domain, and its norm taken as

$$\text{either} \quad \|T_K\|_\infty = \sup_{\omega\in\mathbf{R}} \|T_K(i\omega)\| \qquad \text{or} \qquad \|T_K\|_\infty = \sup_{\theta\in[0,2\pi]} \|T_K(e^{i\theta})\|\,,$$

where $\|T_K(z)\|$ is the matrix norm of $T_K(z)$, i.e. its maximum singular value.

Problem $\mathcal{P}_\gamma$: Given a positive number γ, does there exist an (internally stable) stabilizing controller K such that $\|T_K\|_\infty \le \gamma$, and if yes find (at least) one.

1.3 Equivalence with a game problem

Assume that the disturbance $w(.)$, as a time function, is L^2. (Square integrable in the continuous time case, square summable in the discrete time case). Then, if T_K has a finite H_∞ norm, the output z will also be L^2. Thus, T_K may be viewed as an operator from $(L^2)^\ell$ into $(L^2)^q$. The basic fact is the following one, which easily follows from Parseval's equality :

Proposition 1 *The operator norm of T_K as an operator from $(L^2)^\ell$ into $(L^2)^q$ is the H_∞ norm $\|T_K\|_\infty$ of its transfer function.*

As a consequence, the basic inequality

$$\|T_K\|_\infty \le \gamma \tag{1}$$

is equivalent to

$$\forall w(.) \in (L^2)^\ell, \quad \|z\|_2 \le \gamma \|w\|_2 \,, \tag{2}$$

where $\| \, \|_2$ stands for the L^2 norms, and of course, $z = T_K w$. We shall rather use the equivalent form

$$\forall w(.) \in (L^2)^\ell, \quad \|z\|_2^2 \le \gamma^2 \|w\|_2^2 \,. \tag{2a}$$

This in turn is equivalent to

$$\sup_{w \in L^2} (\|z\|_2^2 - \gamma^2 \|w\|_2^2) \le 0 \,, \tag{3}$$

and the existence of K such that (3) holds is equivalent to

$$\inf_K \sup_{w \in L^2} (\|z\|_2^2 - \gamma^2 \|w\|_2^2) \le 0 \,. \tag{4}$$

Recall that K is the feedback law $u = Ky$, so (4) is basically a game problem, where the information available to the first controller to choose its control $u(t)$ is the set of past values of $y(s)$, $s < t$. If we can solve this game, we can check whether the inf-sup is non positive, and if yes the optimal control law K is a solution of problem $\mathcal{P}_\gamma$. If the inf-sup is positive, possibly $+\infty$, then there is no solution to that problem.

In view of (2), γ will be called an *attenuation factor*. We shall call γ^* the infimum of the attenuation factors that can be achieved.

1.4 State space formulation

As in the classical problem, we shall restrict our attention to the case where the system is finite dimensional. It therefore has an internal representation

$$\sigma x = Ax + Bu + Dw \,, \tag{5a}$$

$$y = Cx + Ew \,, \tag{5b}$$

$$z = Hx + Gu \,. \tag{5c}$$

σx stands for dx/dt in the continuous time case, and $x(t+1)$ in the discrete time case. There is no need to include a term depending on u in y, because it might always be substracted out. And we assume there is no feedthrough from w to z because it simplifies the problem. This corresponds to the classical hypothesis that G_{11} and G_{22} are strictly proper. In the present approach, it is straightforward to raise that restriction, but computationally heavy.

However, contrary to the classical case, we may consider a non stationary problem : the matrices A, B, C, D, E, H and G may be time varying, bounded and piecewise continuous in the continuous time case. We also consider the problem over a finite time interval $[0, T]$. We shall recover the stationary problem by considering the case where the matrices are constant, letting $T = +\infty$

We introduce the notations

$$\begin{pmatrix} H' \\ G' \end{pmatrix} \begin{pmatrix} H & G \end{pmatrix} = \begin{pmatrix} Q & S \\ S' & R \end{pmatrix} \text{ and } \begin{pmatrix} D \\ E \end{pmatrix} \begin{pmatrix} D' & E' \end{pmatrix} = \begin{pmatrix} M & L \\ L' & N \end{pmatrix} \tag{5}$$

We shall always assume the following :

HYPOTHESES

H1. G is injective (one to one)

H2. System (5) is completely observable by z (i.e. equivalently, the pair $[A - BR^{-1}S', (Q - SR^{-1}S')^{1/2}]$ is completely observable) over $[t, T)$, $\forall t \in (0, T)$

H3. E is surjective (onto)

H4. System (5) is completely reachable by w (or more precisely, the pair $[A - LN^{-1}C, (M - LN^{-1}L')^{1/2}]$ is completely reachable) over $[0, t]$, $\forall t \in (0, T]$.

We shall still examine whether there exists a controller such that (2) be satisfied, but with, for any signal $v(.)$

$$\|v\|_2^2 = \int_0^T \|v(t)\|^2 dt \qquad \text{or} \qquad \|v\|_2^2 = \sum_{t=0}^{T-1} \|v(t)\|^2 \tag{6}$$

depending on whether $t \in \mathbf{R}$ or $t \in \mathbf{Z}$.

It will be convenient to further unify the notations of the continuous time and discrete time cases by using the same notations for both cases above, as

$$\|v\|_2^2 = \int_0^T \|v(t)\|^2 dt$$

of course, this will apply to all integrals in the future : the symbol $\displaystyle\int_t^T a(t)\, dt$ will have its classical meaning in the continuous time case, and mean $\displaystyle\sum_{t=0}^{T-1} a(t)$ in the discrete time case.

It also turns out to be natural, in going to the finite horizon case, to add to $\|z\|_2^2$ a penalization on $x(T)$, weighted by some positive (semi) definite matrix X, ending up with the familiar looking performance index

$$J = \|x(T)\|_X^2 + \int_0^T (x', u') \begin{pmatrix} Q & S \\ S' & R \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} dt = \|\zeta\|_2^2\,,$$

where $\zeta = (z(.), x(T)) \in \mathcal{Z} = L^2([0, T] \to \mathbf{R}^q) \times \mathbf{R}^n$.

Dually, since we consider a problem where the controller does not know the state $x(t)$, but only a noise corrupted measurement $y(t)$, it is consistent to consider $x(0) = x_0$ as part of the disturbance. We shall therefore let $\omega = (x_0, w(.)) \in \Omega = \mathbf{R}^n \times L^2([0, T] \to \mathbf{R}^\ell)$, and take as its norm

$$\|\omega\|_2^2 = \int_0^T \|w(t)\|^2 dt + \|x_0\|_Y^2$$

where again, Y is some positive definite matrix, the inverse of which, $Z = Y^{-1}$ will turn out to play a role dual to that of X above.

In the stationary, infinite horizon (classical) case, X and Y will disappear, because, consistent with classical LQ control theory, we shall request that the controller be stabilizing, i.e. $x(t) \to 0$ as $t \to \infty$, and dually we shall let $x(t) \to 0$ as $t \to -\infty$, thus looking at the homogeneous mapping $w(.) \mapsto z(.)$, as does the classical literature which is based upon transfer functions.

Notice also that in the classical literature, one usually makes the simplifying assumption that $S = 0$ and $L = 0$ (i.e., $H'G = 0$ and $DE' = 0$). It is well known that one can achieve that through changes of variables, involving feedback and measurement injection. However, this is a complicated process for a very small simplification, so that we choose here to keep S and L non zero. Just introduce the notations

$$\bar{A} = A - BR^{-1}S', \quad \bar{Q} = Q - SR^{-1}S', \quad \tilde{A} = A - LN^{-1}C, \quad \tilde{M} = M - LN^{-1}L' .$$

Finally, we shall *not* restrict ourselves a priori to linear controllers. Hence, when we shall arrive at the conclusion that the problem has no solution (γ has been chosen too small), this will mean that no controller, either linear or nonlinear, can achieve (2). To recall this, we shall denote by $u = \mu(y)$ the controller. The class $\mathcal{M}$ of admissible controllers will be all causal controllers (depending on the past values of y), such that, when placing $u = \mu(y)$ in the dynamic equation, it has a unique solution $x(.)$ for every square integrable disturbance $w(.)$, leading to a square integrable $z(.)$ (i.e., although possibly non linear, T_μ is still an operator from $(L^2)^\ell$ into $(L^2)^q$).

Thus the problem we consider will be

$$\inf_{\mu\in\mathcal{M}} \sup_{\omega\in\Omega} (\|\zeta\|_2^2 - \gamma^2\|\omega\|_2^2) . \tag{7}$$

The question is to know whether this inf sup is non positive. As a matter of fact, because it is a quadratic form under linear constraints, it will turn out that it is either zero or plus infinity, and when it is zero, a linear map μ will do.

2 Fundamentals of dynamic games

2.1 Information structures

Let first $\mathcal{U}$ and $\mathcal{W}$ be two (decision) sets, and $J : \mathcal{U} \times \mathcal{W} \to \mathbf{R}$ be a performance index over $\mathcal{U} \times \mathcal{W}$. It is well known that one always has

$$\sup_{w\in\mathcal{W}} \inf_{u\in\mathcal{U}} J(u,w) \le \inf_{u\in\mathcal{U}} \sup_{w\in\mathcal{W}} J(u,w).$$

If the equality holds, the game is said to have a *value*. If moreover the inf sup and the sup inf are a min-max and a max-min, then let (u_1, w_1) be an argument of the min max, it is then a saddle point :

$$\forall u \in \mathcal{U},\ \forall w \in \mathcal{W}, \quad J(u_1, w) \le J(u_1, w_1) \le J(u, w_1). \tag{8}$$

Conversely, if a saddle point exists, then min max equals max min. And if (u_i, w_i), $i = 1, 2$ are two saddle points, (u_i, w_j), $j \ne i$ provide two other saddle points yielding the same value.

Further structure arises when $\mathcal{U}$ and $\mathcal{W}$ are sets of time functions, J being given through a first order dynamic equation

$$\sigma(x) = f(t; x, u, w), \quad x(0) = x_, \quad u \in U, \quad w \in W , \tag{9}$$

(again, σx stands for dx/dt or $x(t + 1)$ depending on the framework) and an additive pay-off

$$J = M(x(T)) + \int_0^T L(t; x, u, w)\, dt , \tag{10}$$

with our standing assumption on the meaning of the integral sign holding. Then the important point to discuss is : what information is available to each player at time t to chose its control, $u(t)$ or $w(t)$.

The simpler case is when both players play "open loop", i.e. have no additional information beyond the rules of the game. Then this is essentially a static game. Notice however that in case there is no saddle point, the inf sup, or min max if it exists, is not a very operational concept. Indeed, the first player, if it wants to play "safe" may chose this strategy, say $\hat{u}$. But there is actually no way in which the second player could actually play according to the *same* solution concept, because the argument $\hat{w}$ of the $\max_w J(u,w)$ is itself a function of u, that the second player does not know. It might *anticipate* that its opponent is a safe player, and play $\hat{w} = \arg\max_w J(\hat{u},w)$. But this is not a safe behaviour, because player one might have played otherwise.

This is *not* the case for a saddle point, where relation (9) says that, not only w_1 is optimal against u_1 (this is the left inequality) but also it is safe (this is the right inequality). This, of course, holds symmetrically for u_1.

Assume now that player one is able to play in "closed loop", i.e., at time t, it knows $x(s)$, $s \leq t$, and information about the past values of its opponent's control values $w(s)$, $s < t$. And allow it to use a causal controller $u = \mu(x)$, i.e.

$$u(t) = \mu(t; x(s),\ s \leq t). \tag{11}$$

Then, for a given such controller, one can define $\sup_{w\in\mathcal{W}} J(\mu,w)$. And because the behaviour of the first player is now frozen, this sup does not depend on the information available to compute $w(t)$: we are confronted with a one player deterministic control problem. Therefore, as previously, we may consider the problem

$$\inf_\mu \sup_w J(\mu,w)$$

and the inf sup does not depend on the information assumed to be available to the second player, while it does depend (as we shall see further) on the fact that the first one now plays closed loop.

A further fact is that, because of the "updating theorem" (see [4]), for a finite duration game satisfying, "Isaacs'condition", i.e. such that the hamiltonian has a saddle point in (u,v), it suffices to consider the particular subset of the closed loop strategies called the state feedback strategies, or markovian strategies, of the form

$$u(t) = \mu(t; x(t))\,. \tag{12}$$

If the hamiltonian does not have a saddle point, it may still be possible to define, say, an upper value, replacing the saddle point in (16a) by a min-max, leading to a pair of "upper startegies" where $v(t) = \nu^*(t; x(t), u(t))$ explicitly depends on the current control $u(t)$.

Furthermore, if there exists a unique state feedback saddle point, there are an infinity of closed loop saddle points, but all yield the same state history $x(.)$ and control histories $u(.), w(.)$. Moreover, the state feedback saddle point is the only *strongly consistent* optimal strategy : whatever the past controls, using that strategy from current time to T yields the best guaranteed payoff still possible. This is the essence of the "representation theorem" (see [3]).

2.2 Dynamic programming and Isaac's equation

The reliance on markovian strategies allows one to solve full state information games via dynamic programming, or, in the continuous time case, the Hamilton Jacobi Isaacs equation. We shall restrict our attention to the case where the game (10), (11) has a *closed loop saddle point*. We have stressed that this also yields the closed loop-open loop min max, that we are seeking.

Let $V : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}$ be a function, of class C^1 in the continuous time case, and denote

$$\delta V(t; x, u, w) = V(t+1, f(t; x, u, w)) - V(t, x) \tag{14a}$$

in the discrete time case, and

$$\delta V(t; x, u, w) = \frac{\partial V}{\partial t}(t, x) + \frac{\partial V}{\partial x}(t, x) f(t; x, u, w) \tag{14b}$$

in the continuous time case. And let

$$H(t; x, u, w) = \delta V(t; x; u; w) + L(t; x, u, w). \tag{15}$$

We have the following fundamental result.

Theorem (Isaacs). If there exists a function $V : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}$, and a pair of admissible feedbacks $\mu^* : \mathbf{R} \times \mathbf{R}^n \to U$, $\nu^* : \mathbf{R} \times \mathbf{R}^n \to W$, such that $\forall x, t \in \mathbf{R}^n \times [0, T]$, $\forall u \in U$, $\forall w \in W$,

$$H(t; x, \mu^*(t, x), w) \leq 0 = H(t; x, \mu^*(t, x), \nu^*(t, x)) \leq H(t; x, u, \nu^*(t, x))\,, \tag{16a}$$

$$\forall x \in \mathbf{R}^n, \quad V(T, x) = M(x)\,, \tag{16b}$$

then (μ^*, ν^*) is a closed loop saddle point, and the optimal value of the game is $V(0, x_0)$.

Proof. The proof is entirely straightforward, noticing that summing (or integrating) H from 0 to T on a trajectory yields

$$\int_0^T H(t; x(t), u(t), w(t))\, dt = J(u(.), w(.)) - V(0, x_0).$$

A converse exists in the discrete time case. The discrete Isaacs equation yields a way to compute $V(t, x)$ retrogressively from $V(T, x)$: the dynamic programming algorithm. If at a step t, $\inf_u \sup_w H = +\infty$ for all x, then the problem (10), (11) has itself an infinite $\inf_\mu \sup_w$. This easily follows from the fact that from a state x at time t, the second player can insure himself a payoff at least $V(t, x)$.

The same converse does not hold in the continuous time case, because the value function might not be of class C^1. A specific proof has to be developped for the linear quadratic case, yielding the needed result.

3 Partial information

3.1 The certainty equivalence principle

The problem $\bar{\mathcal{P}}$ of interest to us is one of the form (10) (11), but where the information available to the first player is not the full state, but a measurement sequence of the form

$$y(t) = h(t; x(t), w(t)). \tag{17}$$

To account for the fact that x_0 may be unknown, we also modify the performance index, and let it be

$$\bar{J}(u,w) = M(x(T)) + \int_0^T L(t; x(t), u(t), w(t))\, dt - N(x_0)\,.$$

It turns out that, if the full state information game (10) (11) has a (state feedback) saddle point (μ^*, ν^*), then under suitable conditions, the min max of this partial information game can be obtained via a *certainty equivalence principle* : one can compute recursively, from the available information, an "estimate" $\hat{x}(t) \in \mathbf{R}^n$ such that the strategy

$$\hat{\mu}(y)(t) = \mu^*(t, \hat{x}(t)) \tag{18}$$

is a min max strategy.

The estimate $\hat{x}$ is obtained by solving, at each instant of time, an auxiliary problem, that we now define. Let $\tau \in (0,T)$. Let, for any time function $v(.)$, v^τ denote the set $\{v(t),\ t < \tau\}$. To a given pair (u^τ, y^τ), associate the set Ω_τ of disturbances $(x_0, w(.))$ (or equivalently ω) such that, when placed together with u^τ in (10) and (17), they generate the given y^τ. We write this in short

$$\Omega_\tau(u^\tau, y^\tau) = \{\omega \mid u^\tau, y^\tau\} \tag{19}$$

and call it the set of disturbances *compatible* with (u^τ, y^τ). Next define an auxiliary performance index

$$G^\tau(u^\tau, \omega^\tau) = V(\tau, x(\tau)) + \int_0^\tau L(t; x(t), u(t), w(t))\, dt - N(x_0) \tag{20}$$

and an auxiliary problem :

Problem $\mathcal{Q}^\tau(u^\tau, y^\tau)$:

$$\max_{\omega^\tau \in \Omega_\tau(u^\tau, y^\tau)} G^\tau(u^\tau, \omega^\tau)\,.$$

Then, if $\mathcal{Q}^\tau$ has a unique solution, say $\hat{\omega}^\tau(.)$ generating a trajectory $\hat{x}^\tau(.)$, let $\hat{x}^\tau(\tau) = \hat{x}(\tau)$ for short. This is our "estimate".

We have the following theorem ([6], [2]).

Theorem 1 *(min max certainty equivalence principle). If the full state information game (10) (11) has a unique saddle point (μ^*, ν^*) leading to a (C^1 in the continuous time case) value function $V(t,x)$, and if all auxiliary problems $\mathcal{Q}^\tau$ have unique solutions, (which are the only solutions of the variational inequalities ; this extra condition being needed only in the discrete time case) then the strategy (18), where $\hat{x}(\tau)$ is the state at time τ in the maximizing trajectory of $\mathcal{Q}^\tau(u^\tau, y^\tau)$, is a min-max strategy for the problem $\bar{\mathcal{P}}$. Furthermore, the value of the min-max is the same as that of the full information saddle point.*

The proof is a bit too technical to be given in the framework of this short lecture. As a matter of fact, it gives a slightly more powerful result : whatever u^τ, use of $u(t) = \hat{\mu}(y)(t)$ from time τ on insures the best possible guaranteed pay off given the available information at time τ. In that sense, $\hat{\mu}$ is *strongly consistent*. It can be extended to the case where (μ^*, ν^*) is only an upper saddle point.

3.2 Forward-backward dynamic programming

The above theory seems difficult to apply, since at each instant of time $\tau \in (0, T)$ a new constrained maximization problem $\mathcal{Q}^\tau$ must be solved to get $\hat{x}(\tau)$. However, dynamic programming can again be used to provide a recursive computation of $\hat{x}$.

Let $W(t,x)$ be a function from $\mathbf{R} \times \mathbf{R}^n$ into $\mathbf{R}$, and construct

$$K(t; x, u, w) = \delta W + L \tag{21}$$

as in (14) (15), with V replaced by W. Let also Ψ_t be a set of instantanenous disturbances given by

$$\Psi_t(x, y) = \{w \mid h(t; x, w) = y\}\,. \tag{22}$$

Assume that for a given pair of time functions $\{u^\tau(.), y^\tau(.)\}$, one has

$$\forall (t,x) \in [0,\tau) \times \mathbf{R}^n, \quad \exists \hat{w} \in \Psi_t(x, y(t)) : \quad \forall w \in \Psi_t(x, y(t)),$$

$$K(t; x, u(t), w) \leq 0 = K(t; x, u(t), \hat{w})\,, \tag{23a}$$

$$\forall x \in \mathbf{R}^n, \quad W(0,x) = N(x)\,. \tag{23b}$$

Then, whatever x_0 and the sequence w^τ, provided they be compatible with (u^τ, y^τ), i.e. $w(t) \in \Psi_t(x(t), y(t))$, $\forall t \in [0, \tau)$, we have by summing (23a) and using (23b), as in Isaac's theorem

$$G^\tau(u^\tau, w^\tau) \leq V(\tau, x(\tau)) - W(\tau, x(\tau))\,,$$

with the equality possible if $w(t) = \hat{w}(t)$, $\forall t < \tau$.

Assume finally that all $x \in \mathbf{R}^n$ are reachable at time τ by maximizing trajectories. This assumption is natural since x_0 is free. Then the auxiliary problem $\mathcal{Q}^\tau(u^\tau, y^\tau)$ has a finite maximum if and only if there exists

$$\max_{x \in \mathbf{R}^n} (V(\tau, x) - W(\tau, x)) = V(\tau, \hat{x}(\tau)) - W(\tau, \hat{x}(\tau)). \tag{24}$$

Notice that (23) provides a way of computing $W(t,x)$ from $W(0,x) = N(x)$:

$$\frac{\partial W}{\partial t}(t,x) = -\max_{w \in \Psi_t(x, y(t))} \left(\frac{\partial W}{\partial x} f(t; x, u(t), w) + L(t; x, u(t), x) \right) \tag{25}$$

in the continuous time case and a slightly more complicated dynamic programming algorithm in the discrete time case : let

$$\Phi_t = \{(w, \xi) \in W \times \mathbf{R}^n \mid f(t; \xi, u(t), w) = x, \quad h(t; \xi, w) = y(t)\}\,,$$

the associated algorithm is

$$V(t+1, x) = -\max_{(w,\xi) \in \Phi_t} (-V(t, \xi) + L(t; \xi, u(t), w)). \tag{26}$$

4 Finite horizon minimax (H_∞ bounded) controller

We now examine the problem (8) for the possibly non stationary system (5). We first need to find the state feedback saddle point, and its condition of existence. This obliges us to deal separately with the continuous time and the discrete time problems.

4.1 Continuous time case

Let us first look at the state feedback problem. It is a well known fact that in the Linear-Quadratic (LQ) problem, one can seek a value function of the form

$$V(t,x) = \|x\|^2_{P(t)}, \tag{27}$$

and the Hamilton-Jacobi Isaacs equation degenerates into the celebrated Riccati equation

$$\dot{P} + P\bar{A} + \bar{A}'P - P(BR^{-1}B^1 - \gamma^{-2}DD')P + \bar{Q} = 0\,, \tag{28a}$$

$$P(T) = X\,, \tag{28b}$$

and the saddle point strategies are given, if $P(.)$ exists by

$$\mu^*(t,x) = -R^{-1}(B'P + S')x, \tag{29a}$$

$$\nu^*(t,x) = \gamma^{-2}D'Px. \tag{29b}$$

Existence of $P(.)$ solution of (28), over $[0,T]$ is a sufficient condition for the existence of a saddle point. The reciprocal is a bit more complex. If the Riccati equation fails to have a solution, i.e. the matrix $P(.)$, when integrated backward from T, diverges before t reaches 0, we say it has a *conjugate point* in $[0,T]$. The game may still have a value, and a saddle point in some sense (see [5]). But this is a limiting case, as the following result shows ([2], chap. 8).

Theorem 2 *If the Riccati equation (28) has a solution over* $[0,T]$*, then the state feedback differential game has a saddle point, given by (29), and the value of the game is given by (27). Conversely, if the Riccati equation has a conjugate point in* $[0,T]$*, then for any smaller* γ *the game has an infinite inf sup.*

Let us now now apply the forward dynamic programming of section 3.2 to find $\hat{x}$ to be used in (29) according to the certainty equivalence principle. Thus we want to solve the p.d.e. (25). Again, we try a quadratic function W, but we have to take it nonhomogeneous :

$$W(t,x) = \|x - \check{x}(t)\|^2_{K(t)} + k(t) \tag{30}$$

where we have to choose $\check{x}(t) \in \mathbf{R}^n$, the symmetric matrix $K(t)$, and the scalar function $k(t)$.

The constraint Ω_t, or Ψ_t in (22), is affine. The constrained maximization in (25) is easily performed via a Lagrange multiplier method. Identification of the quadratic terms in x to zero yields a differential equation for K, identification of the linear terms one for $\check{x}$ and of the constant terms (w.r.t. x) one for k. One easily gets

$$\dot{K} = -K\tilde{A} - \tilde{A}'K + \gamma^2 C'N^{-1}C - Q - \gamma^{-2}K\tilde{M}K\,, \tag{31a}$$

$$K(0) = \gamma^2 Y\,. \tag{31b}$$

It turns out to be convenient to write everything in terms of

$$\Sigma(t) = \gamma^2 K^{-1}(t) \tag{32}$$

which, in view of (31) satisfies

$$\dot{\Sigma} = \tilde{A}\Sigma + \Sigma\tilde{A}' - \Sigma(C'N^{-1}C - \gamma^{-2}Q)\Sigma + \tilde{M}\,, \tag{33a}$$

$$\Sigma(0) = Y^{-1} = Z. \tag{33b}$$

Similarly, and substituting (32) in the equation, it comes

$$\begin{aligned} &\dot{\check{x}} = A\check{x} + Bu + \gamma^{-2}\Sigma H'(H\check{x} + Gu) + (\Sigma C' + L)N^{-1}(y - C\check{x}) \,, \\ &\check{x}(0) = 0 \,. \end{aligned} \tag{34}$$

One also gets an expression for $\dot{k}$, but we do not need to write it.

Finally, (24) becomes here

$$\max_x \left[\|x\|^2_{P(t)} - \|x - \check{x}(t)\|^2_{K(t)} \right] \,.$$

Hence the problem is strictly concave if and only if $P - K < 0$, which we may write in terms of the spectral radius $\rho(\Sigma P)$

$$\forall t \in [0, T], \quad \rho(\Sigma(t)P(t)) < \gamma^2 \tag{35}$$

and leads to

$$\hat{x}(t) = \Big(I - \gamma^{-2}\Sigma(t)P(t)\Big)^{-1} \check{x}(t). \tag{36}$$

The dynamic programming approach followed here shows that if both Riccati equations (28) and (33) have solutions over $[0, T]$, that satisfy (35), then all problems Q^τ being strictly concave, the problem $\bar{\mathcal{P}}$ has a solution, obtained by applying (18), i.e. here

$$u(t) = -R^{-1}(B'P + S')\hat{x}(t) \,. \tag{37}$$

The converse is slightly more technical to prove. One however obtains the following result.

Theorem 3 *If both Riccati equations (28) and (33) have solutions over $[0, T]$, that satisfy (35), then $\gamma \geq \gamma^*$, and a controller that guarantees an attenuation level γ is given by (37), where $\hat{x}$ is given by (34) (36). If one of the three conditions fail, then $\gamma \leq \gamma^*$ (i.e. for any smaller attenuation level, the problem has no solution).*

A remarkable fact is that the system (34)(36) admits an alternate representation, as can be checked by direct calculations:

$$\dot{\hat{x}}(t) = A\hat{x} + B\mu^*(\hat{x}) + D\nu^*(\hat{x}) + \Big(I - \gamma^{-2}\Sigma P\Big)^{-1} (\Sigma C' + L)N^{-1}(y - \hat{y})$$

where $\hat{y} = C\hat{x} + E\nu^*(\hat{x})$, as expected. This looks very much indeed like a standard Kalman filter. Moreover, multiplying this differential equation to the left by $(I - \gamma^{-2}\Sigma P)$ allows one to write a complete solution to the problem without ever inverting this matrix. This is usefull if one wants to use the minimum γ^* for which the problem still has a solution. As a matter of fact, under fairly general conditions, this will be characterized by the fact that this matrix fails to be invertible. Then we still have a controller, although the "filter" equation for $\hat{x}$ is now an *implicit* differential equation. One then needs the theory of implicit systems.

4.2 Discrete time case

The discrete dynamic programming equation (16) for the full information (state feedback) problem again has a solution of the form (27), where P_t is given by the discrete Riccati equation

$$P_t = \bar{A}' P_{t+1} \bar{A} - \bar{A}' P_{t+1} [B \quad D] \begin{pmatrix} R + B' P_{t+1} B & B' P_{t+1} D \\ D' P_{t+1} B & -\gamma^2 I + D' P_{t+1} D \end{pmatrix}^{-1} \begin{bmatrix} B' \\ D' \end{bmatrix} P_{t+1} \bar{A} + \bar{Q}\,, \quad (38a)$$

$$P_T = X\,. \quad (38b)$$

Because of the hypothesis H2, whenever P_{t+1} exists, it is positive definite and (38a) may be written alternatively

$$P_t = \bar{A}^{-1} (P_{t+1}^{-1} + BR^{-1}B' - \gamma^{-2} DD')^{-1} \bar{A} + \bar{Q} \quad (38c)$$

The maximization problem in the dynamic programming is strictly concave if and only if

$$\gamma^2 I - D' P_{t+1} D > 0 \quad (39)$$

which can easily be shown to be equivalent to $P_{t+1}^{-1} - \gamma^{-2} DD' > 0$, so that the inverse in (38c) a fortiori exists. If $\gamma^2 I - D' P_{t+1} D$ has a negative eigenvalue, the supremum in w is always infinite, and so is the inf sup of the game. If it has a zero eigenvalue, it is easy to see that for any smaller γ it would fail to be nonnegative definite. We therefore end up with the following theorem.

Theorem 4 *If the Riccati equation (38) has a solution that satisfies (39) for all $t \in [0, T-1]$ then the full information game has a unique state feedback saddle point*

$$\begin{aligned} \mu^*(t,x) &= -R^{-1}[B'(P_{t+1}^{-1} + BR^{-1}B' - \gamma^{-2}DD')^{-1}\bar{A} + S']x, \\ \nu^*(t,x) &= \gamma^{-2}[D'(P_{t+1}^{-1} + BR^{-1}B' - \gamma^{-2}DD')^{-1}\bar{A} + S']x. \end{aligned} \quad (40)$$

If for some t, (39) fails to hold, then $\gamma \le \gamma^*$: for any smaller γ the game has an infinite supremum in w for any strategy μ.

We turn now to the imperfect information problem. The forward dynamic programming algorithm again has a solution of the form (30). The detailed theory is a little more complicated than in the continuous time case. One has to notice that for Q^τ to be strictly concave, K has to be positive definite. In any extent, one arrives at the equations, dual to the previous one :

$$\Sigma_{t+1} = \tilde{A} \Sigma_t \tilde{A}' - \tilde{A} \Sigma_t [C' \quad H'] \begin{pmatrix} N + C\Sigma_t C' & C\Sigma_t H' \\ H\Sigma_t C' & -\gamma^2 I + H\Sigma_t H' \end{pmatrix} \begin{bmatrix} C \\ H \end{bmatrix} \Sigma_t \tilde{A}' + \tilde{M}\,, \quad (41a)$$

$$\Sigma_0 = Y^{-1} = Z\,, \quad (41b)$$

and because of H4, (41a) takes the alternative form

$$\Sigma_{t+1} = \tilde{A} (\Sigma_t^{-1} + C' N^{-1} C - \gamma^{-2} Q)^{-1} \tilde{A}' + \tilde{M}. \quad (41c)$$

Now, dynamic programming leads to a strictly concave problem if and only if

$$\Delta_t = \Sigma_t^{-1} + C' N^{-1} C - \gamma^{-2} Q > 0, \quad (42)$$

and one may as previously argue that if it is only nonnegative definite at some time t, for any smaller γ it would fail to be so at some time τ, and the problem Q^τ would have an infinite supremum. We also obtain

$$\check{x}_{t+1} = A\check{x}_t + Bu + \gamma^{-2}\tilde{A}\Delta_t^{-1}H'(H\check{x}_t + Gu) + (\tilde{A}\Delta^{-1}C' + L)N^{-1}(y_t - C\check{x}_t)\,, \tag{43a}$$

$$\check{x}_0 = 0\,, \tag{43b}$$

and finally, the problem Q^τ is globally concave as previously if and only if furthermore (35) holds, and $\hat{x}$ will again be given by (36). So that we get the following complete theorem.

Theorem 5 *If the solutions of the Riccati equations (38) and (41) satisfy the conditions (39), (42) over* $[0, T-1]$ *and (35), then* $\gamma \geq \gamma^*$*, and an optimal controller, that guarantees an attenuation level* γ*, is obtained by computing* $\hat{x}$ *via (43) and (36), and substituting it to* x *in (40). If one of the three conditions fails at some time, then* $\gamma \leq \gamma^*$ *: for any lower* γ *the problem has no solution.*

We are not aware at this time of an equivalent to the alternate representation.

We have a rather complete answer for the finite time problem. In both the continuous time and discrete time problems, we have two dual Riccati equations with a condition on each of them, the *conjugate point conditions* and the global concavity condition $\rho\left(\Sigma(t)P(t)\right) < \gamma^2$, sometimes called the *spectral radius condition*. If all conditions are satisfied, then we have an n-dimensional controller that solves the "H_∞-bounded" problem. If one fails, then γ is at best the limiting one γ^* or is smaller than γ^*. It should be emphasized, however, that this solution is highly non unique. It has been part of the classical problem to characterize *all* (linear) controllers that satisfy the H_∞ norm bound. The present approach may not be the best suited to that end. See however in [2] the derivation of a very large class of possibly nonlinear variants in the continuous time case.

Some more analysis must be performed to show what happens if we assume that x_0 is known to be zero. The result holds good under the added hypothesis that (A, D) is completely reachable over $(0, t)$, $\forall t \in (0, T)$, just taking $Z = 0$ to initialize the second Riccati equation. Intuitively, this amounts to placing an "infinite weight" on x_0 in the norm of the disturbance, hence constraining it to be zero.

5 The stationary problem

We are now in a position to turn back to the original stationary problem. It turns out that in the linear quadratic optimization theory that we are extensively using, going from the finite horizon problem to the stationary infinite horizon one is simpler when one furthermore restricts the allowable controllers to those that stabilize the system. However this requirement is precisely part of the classical H_∞ problem. As a consequence the above analysis will carry over very easily.

Let us first consider the full state information problem, which is a perequisite. The problem considered is defined by (5), with constant matrices and (with a continuous or discrete sum)

$$J = \int_0^\infty \|z(t)\|^2 dt. \tag{44}$$

We need to consider the Riccati equation for $P(t)$, either (28a) or (38a), integrated backward in time from $P(0) = 0$. If it exists, call P^* the limit of its solution $P(t)$ as $t \to -\infty$. Then under our hypothesese, it is positive definite, reached from below, it is the maximum solution of the algebraic Riccati equation obtained by letting $P(t) =$ constant in the equation. And furthermore, the closed loop system

$$\sigma x = Ax + B\mu^*(x)$$

where $P(t)$ has been replaced by P^* in μ^*, is asymptotically stable.

The following theorem was proved by Mageirou [9] (see also [3]).

Theorem 6 *The criterion (44) has a inf sup less than $+\infty$ if and only if P^* exists, and in the discrete time case satisfies (39). In that case, substituting P^* for $P(t)$ in μ^* as given by (29) or (40) yields the min sup.*

We now turn to the imperfect information problem. The certainty equivalence problem holds, with

$$G^0 = V(0, x) + \int_{-\infty}^{0} L(x, u, w)dt$$

and imposing that $x(t) \to 0$ as $t \to -\infty$. The system obtained by maximizing G^0 w.r.t. w (over the disturbances compatible with the current information) is stable, thus unstable as $t \to -\infty$, so that the constraint uniquely defines $\hat{x}(0)$.

The maximization of G^0 still involves the Riccati equation (31a), which through the transformation (32) yields the equation (33) in Σ, or its counterpart (41) in the discrete time case. Again, integrated forward from $\Sigma(0) = 0$, it may have a limit Σ^* which is then, under our hypothesese, positive definite, reached from below, it is the maximum solution of the related algebraic Riccati equation, and stabilizes the system (34) or (43).

The global concavity condition is unchanged, and must hold with Σ^* and P^* :

$$\rho(\Sigma^* P^*) < \gamma^2. \tag{45}$$

Altogether, this yields the following theorem.

Theorem 7 *If both Riccati equations admit limit values P^* (satisfying (39) in the discrte time case) and Σ^* (satisfying (42) in the discrete time case), jointly satisfying (45), then replacing $P(t)$ and $\Sigma(t)$ by P^* and Σ^* in the equations of the (finite horizon) controller yields a stable, stabilizing stationary controller insuring an attenuation level γ. If one of the above conditions fails, $\gamma \leq \gamma^*$.*

It is worthwhile mentioning the fact that P^* and Σ^* can be characterized in terms of the stable subspaces of $2n \times 2n$ hamiltonian matrices, and that this yields another way of computing them.

6 Concluding remarks

The solution of the basic problem (1) is usally non unique, particularily if γ is not the minimum possible one γ^*. The one we have given here is known in the literature as the "central solution".

It is a very remarkable fact that it resembles so much the classical LQG regulator. In essence, the value function associated to our auxiliary problem, a “conditional maximization” problem, plays the role of the conditional probability law in stochastic control. As a matter of fact, the classical LQG regulator is obtained as the formal limit of this “central” regulator as $\gamma \to \infty$.

This together with the state space formulation of the problem, suggests a slightly different motivation for the new theory, beyond the “classical” motivations, as related in [8], [9] for instance. All pertain to some form of robustness.

Let us take it for granted that there is some merit to the following way to derive a controller for a dynamical system : fit a linear model to your data on the system, choose a quadratic performance index, J, and choose a control that “keep it small in spite of the disturbances”. Now, these disturbances do exist, otherwise there would be no control problem. They come from the fact that the real system may deviate from its linear approximation, and from physical disturbances either in the dynamics (wind, and other variations on “natural” phenomena) or as measurement noise. The problem, thus, is to find a way to devise a rational behaviour in presence of these unknown disturbances.

The classical approach, stochastic control, is in a large extent inspired by the experience of electric noise in electric devices, turbulence in gases, and the like. These are high frequency phenomena, with a predictible average value in the long run. Hence the idea to represent them as stochastic processes, white noises of known covariance, appealing to the theory of probabilities. Once it is decided to do so, a rational behaviour is to minimize the expectation of J, hence the LQG controller.

Notice, however, that how compelling they may sometimes appear, probabilities are not in the physical world : they are our mental construct to model it, and it is our choice to use them in the control problem at hand.

While some types of distrubances may fit well the white noise model (or a process with Gauss-Markov representation, modelled through a shaping filter included in the plant model), others may not. This may be the case if the bulk of these unknown quantities comes from poor modelling, either because the plant is very nonlinear, or because it is ill known. We argue that in that case, it may be more efficient ot use a different approach, namely the one we explain in section 1.2 above. It is plain that if w is “large” we wont be able to keep J small. Let us at least try to control J's rate of growth, i.e. $J/\|w\|^2$. This is what the new theory does.

There remains to accumulate more engineering experience to know whether it behaves better, when, and to what extent.

References

[1] T. Başar: “Game theory and H_∞-optimal control”, *4th international conference on Differential Games and Applications*, august 1990, Springer Verlag.

[2] T. Başar and P. Bernhard: *H_∞-Optimal Control and Related Minimax Design Problems, A Dynamic Game Approach*, Birkhauser, 1991

[3] T. Başar and G.J. Olsder: *Dynamic Noncooperative Game Theory*, Academic Press, 1982

[4] P. Bernhard: “Information necessaire pour la commande dans le cas le plus defavorable” *Rapport CEREMADE 7923*, Université Paris 9 Dauphine, 1979

[5] P. Bernhard: "Linear-quadratic two-person zero-sum differential games, necessary and sufficient conditions", *JOTA*, **27** pp 51–69, 1979

[6] P. Bernhard: "A min-max certainty equivalence principle and its applications to continuous time, sampled data, and discrete time H_∞ optimal control", *INRIA report 1347*, august 1990

[7] J. Doyle, K. Glover, P. Khargonekar, and B. Francis: "State-space solutions to standard H_2 and H_∞ control problems", *IEEE Transactions in Automatic Control*, **AC-34**(8), pp 831–847, 1989

[8] B. Francis: *A course in H_∞ control theory*, Lecture notes in control and information sciences, **88**, Springer-Verlag, 1987

[9] H. Kwakernaak: "The Polynomial Approach to H_∞-optimal Regulation", *Lecture Notes 1990 CIME Course on Recent Developments in H_∞ Control Theory*, Villa Olmo, June 18-26, 1990, Italy. To be published by Springer Verlag.

[10] E. F. Mageirou: "Value and strategies for infinite duration linear quadratic games", *IEEE Transactions on Automatic Control*, **AC 21**(4), pp 547–550, 1976

[11] G. Zames: "Feedback and optimal sensitivity: Model reference transformation, multiplicative seminorms and approximate inverses", *IEEE Transactions on Automatic Control*, **AC 26**, pp 301–320, 1981