Max-plus algebra and mathematical fear in dynamic optimization

Pierre Bernhard*

October 1999

Abstract

Max plus algebra, cost measures, and mathematical fear have proved usefull tools in dynamic optimization. Indeed, the first two have even become a central tool in some fields of investigation such as discrete event systems. We first recall the fundamentals of max plus algebra with simple examples of max-plus linear models, and simple consequences of that remark. We then introduce cost measures, the natural equivalent of probability measures in the max-plus algebra, and their fundametal properties, including the definition of the mathematical fear (the equivalent of the mathematical expectation), induced measures and conditioning. Finally, we concentrate on those aspects that are put in use in dynamical optimization and state a separation theorem which was first derived using these tools.

1 Introduction

The max plus algebra is just a special case of Maslov's idempotent algebras, exactly as cost measures are a special case of his idempotent measures. It was the merit of Quadrat to stress the importance of this special case and to develop with his coworkers a beautiful theory. Rather than attempt an exhaustive bibliography here, we have rather refer the reader to the magnificent book [2], from which the begining of this presentation is inspired.

We have used that aparatus, stressing the role of mathematical fear[1], a concept parallel to mathematical expectation, to derive results in partial information min-sup dynamic optimization problems with partial information. It was implicitly at the basis of our derivation of the minimax certainty

*I3S, University of Nice-Sophia Antipolis and CNRS

[1]this phrase is ours

equivalence theorem, and explicitly in the more recent theory of L^1/L^∞ control that we sketch here. We shall therefore develop more specifically the tools that lead to that result.

Notice that extending the parallel with mathematical expectation to the *minimization* of such quantities is made possible by the simple remark of the subsection "mathematical fear" beneath, that if $\varphi(\omega) \le \psi(\omega)$ for all ω, then $\mathbb{F}\varphi \le \mathbb{F}\psi$. An almost trivial remark, but not completely naïve though.

Simultaneously with the development of our theory, James and Baras introduced their "informational state", the exact equivalent of our "cost to go" in [3]. This informational state is an unnormalized conditional state cost measure, easier to calculate than the normalized one, exactly as Zakai's equation is simpler than the non linear filter equation. We showed in [5] that using the "true" (normalized) conditional state cost measure leads to more appealing formulas downstream, as the last section shows.

2 Max plus algebra and linearity

2.1 Max-plus algebra

Consider the set $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty\}$ endowed with the two operations $(\max, +)$. It shares a certain number of resemblances with the "classical" algebra $(+, \times)$ on $\mathbb{R}$. To make them more apparent, we shall for a while write $a \oplus b$ for $\max\{a, b\}$ and $a \otimes b$ for $a + b$. Simultaneously, we shall write e for 0 and ε for $-\infty$. Notice then the following properties :

Property	classical	max-plus
Associativity	$(a+b)+c = a+(b+c)$	$(a \oplus b) \oplus c = a \oplus (b \oplus c)$
Neutral el.	$a + 0 = a$	$a \oplus \varepsilon = a$
Associativity	$(a \times b) \times c = a \times (b \times c)$	$(a \otimes b) \otimes c = a \otimes (b \otimes c)$
Neutral el.	$a \times 1 = a$	$a \otimes e = a$
Inverse	$a \times (1/a) = 1$	$a \otimes (-a) = e$
Distributivity	$a \times (b+c) = (a \times b) + (a \times c)$	$a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c)$
Absorbption	$a \times 0 = 0$	$a \otimes \varepsilon = \varepsilon$

Of course, some properties do not carry over so nicely. For one thing the first operation, max, has no inverse in that new algebra. Thus we do not have a ring, or a fortiori a field, but an algebraic structure sometimes called a *dioide* or *semifield*. Also, we have that for all a, $a \oplus a = a$, hence the name of *idempotent* algebra.

We can of course carry out vector and matrix calculations in that algebra. If $A = \{a_{ij}\}$, $i = 1, \dots, m$ and $j = 1, \dots, n$ is an $m \times n$ matrix and $x = \{x_j\}$,

$j = 1, \ldots, n$ a vector, then $y = A \otimes x$ in that algebra means, in classical notations, that

$$y_i = \max_j \{a_{ij} + x_j\}$$

and it is a simple matter to write a matrix product.

An eigenvalue in that algebra is a number λ such that there exists a vector x with coordinates not all $-\infty$ satisfying

$$A \otimes x = \lambda \otimes x$$

i.e., componentwise and in classical notations

$$\max_j \{a_{ij} + x_j\} = \lambda + x_i\,, \qquad i = 1, \ldots n\,.$$

An important tool in the analysis of square matrices is their precedence graph. It is a theorem that the largest mean weight of any cycle of its precedence graph is an eigenvalue of a square matrix, and the only one if this graph is strongly connected. There also is, in the new algebra, a characteristic equation and a Cayley-Hamilton theorem.

Many applications lead to equations like $x = A \otimes x \oplus b$, that we shall write $x = Ax \oplus b$, ignoring the multiplication sign as in classical algebra. If all eigenvalues of A satisfy $\lambda \le e$ —there is no cycle of positive weight in its graph— it is possible to give a precise meaning to the classical series expansion of $(I - A)^{-1}$, thus definig a linear operator A^* such that the previous equation has $x = A^* b$ as its solution.

2.2 Examples

2.2.1 Discrete events systems

Consider a production scheduling problem. The classical tool to describe the system is that of Petri nets. This paragraph is intended to a reader somewhat familiar with such a description. It can be skipped with no harm for the rest of this paper. However this is currently the main application of the max plus algebra. We find it worthwhile to give a short glimpse at it.

We shall limit ourselves to the special case where the system is described by an event graph, i.e. a Petri net where there is only one transition upstream and downstream of any place. More precisely, to describe a scheduling problem, we need a timed event graph, where each place is associated with a positive number : the time a token must stay in it before it is available to fire the next transition.

Consider the event graph[2] of figure 1. Here, v and w are two inputs

[2]The tradition of Petri nets is to draw nice curly arcs.

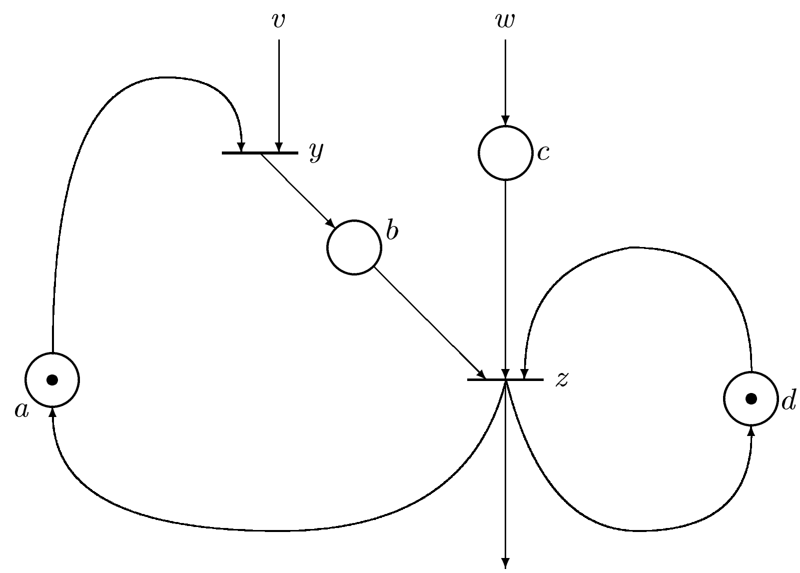

Figure 1: A production scheduling timed event graph

(parts coming into the shop), v_k and w_k are the epoch at which parts enter the graph for the k-th time. Similarly, the letters y and z denote the two transitions of this graph, meaning that y_k and z_k denote the epoch at which their respective firings number k occur. The positive numbers a, b, c, and d are the delays induced in the places they mark.

It is a simple matter to write equations for y_{k+1} and z_{k+1} :

$$\begin{aligned} y_{k+1} &= \max\{a + z_k, v_{k+1}\}\,, \\ z_{k+1} &= \max\{b + y_{k+1}, d + z_k, c + w_{k+1}\}\,. \end{aligned}$$

In the notations of the new algebra, this can be written in terms of the vectors

$$x = \begin{pmatrix} y \\ z \end{pmatrix}\,, \qquad u = \begin{pmatrix} v \\ w \end{pmatrix}\,,$$

and the matrices

$$A_0 = \begin{pmatrix} \varepsilon & \varepsilon \\ b & \varepsilon \end{pmatrix}\,, \quad A = \begin{pmatrix} \varepsilon & a \\ \varepsilon & d \end{pmatrix}\,, \quad B = \begin{pmatrix} e & \varepsilon \\ \varepsilon & c \end{pmatrix}\,,$$

as

$$x_{k+1} = A_0 x_{k+1} \oplus A x_k \oplus B u_{k+1}\,.$$

Hence we have an implicit linear system, and provided that A_0 has nonpositive eigenvalues, this can be transformed into an explicit one

$$x_{k+1} = \bar{A}x_k \oplus \bar{B}u_{k+1}\,,$$

with $\bar{A} = A_0^* A$ and $\bar{B} = A_0^* B$.

Thus we have a linear system. A new timed event graph can be associated with this new formulation. One can deduce from the matrices of the system its asymptotic throughput, its stability, stablize it if it is not stable, etc.

2.2.2 Dynamic Programming

Let X be a finite *state space*, of cardinal $|X| = N$, $x \in X$ the state of a controlled machine, for any $y, z \in X$, let $L(y, z)$ the cost of transitioning from state y to state z, (hence $L : X \times X \to \mathbb{R}$) and $K : X \to \mathbb{R}$ be a final cost, so that the cost of a trajectory $\{x_k\}, k = 0, ..., T$ is

$$J(\{x_k\}) = K(x_T) + \sum_{k=1}^{T} L(x_{k-1}, x_k)\,.$$

We wish to *maximize* the cost for a given initial state. Let $V_T(x_0)$ be the return function of this T step dynamic decision problem with initial state x_0. The dynamic programming equation is

$$\begin{aligned} V_k(x) &= \max_{y \in X}[L(x, y) + V_{k-1}(y)]\,, \\ V_0(x) &= K(x)\,. \end{aligned}$$

Let V_k be the vector of $\mathbb{R}^N$ whose coordinates are the $V_k(x), x \in X$ and let L be the matrix of the $L(x, y)$, $x, y \in X$. In our new algebra, the dynamic programming equation becomes

$$V_k = LV_{k-1}\,, \quad V_0 = K$$

so that this system has an obvious “explicit” solution $V_N = L^N K$, provided that the matrix products be understood in the sense highlighted above.

We shall see below that Quadrat [9] has drawn on that remark and the next section to derive new results concerning the asymptotic behaviour of dynamic programming.

3 Cost measures

3.1 Axioms

Maslov introduced measures in idempotent algebras which were specialized to cost measure, the equivalent of probability measures, by Quadrat and co-workers. Cost measures are functions of subsets of a set Ω. We show here how the axioms of cost measures follow naturaly from those of probability measures by substituting $(\max, +)$, (or equivalently $(\oplus, \otimes)$, but we choose to switch back to traditional notations, leaving it to the reader to convince himself that this is indeed the natural parallel) to $(+, \times)$. We write $p(A)$ and $c(A)$ respectively for the probability or the cost of a subset $A \subset \Omega$, with A belonging to a given sigma algebra in the first case, to the borelian of a topology in the second case.

Property	Probability	Cost
Empty set	$p(\emptyset) = 0$	$c(\emptyset) = -\infty$
Disjoint sets	$p(A \cup B) = p(A) + p(B)$	$c(A \cup B) = \max\{c(A), c(B)\}$
Normalisation	$p(\Omega) = 1$	$c(\Omega) = 0$
Bayes' rule	$p(A\|B) = p(A \cup B)/p(B)$	$c(A\|B) = c(A \cup B) - c(B)$

One better understands the axioms of the cost measures once densities are introduced. Let p have a density P and c have a density Q. (It is a theorem of Mariane Akian [1] that in some sense, every cost measure has a density. See also [4].) Then we have

$$p(A) = \int_{\omega \in A} P(\omega)\, d\omega\,, \qquad c(A) = \sup_{\omega \in A} Q(\omega)\,.$$

This last definition induces the first two properties of the above table. Together with the third one, it implies that $Q(\omega)$ is nonpositive for all $\omega \in \Omega$.

There would be much more to say on conditioning. Notice that the knowledge that $\omega \in B$ naturaly leads to the operation $\max_{\omega \in A \cap B} Q(\omega)$, which is $c(A \cap B)$. The extra term $-c(B)$ is exactly what is needed to normalize the conditional cost distribution.

Two events A and B are said to be *independent* if $c(A|B) = c(A)$, which in view of Bayes'rule leads, as one would expect, to $c(A \cap B) = c(A) + c(B)$. As a consequence decision variables (see below) will be called *independant* if their joint cost distribution is the sum of their individual (marginal) cost distributions.

3.2 Decision variables

3.2.1 Probability distribution

A continuous real function over Ω is called a *decision variable*. It is associated with a cost measure and a cost probability in the natural way : if $x = X(\omega)$ is a decision variable, its cost density can be obtained as

$$Q_X(x) = c(X^{-1}(x)) = \sup_{\omega | X(\omega)=x} Q(\omega) . \tag{1}$$

This same technique propagates an induced cost density on functions of x : if $y = h(x)$, (h continuous) it is a decision variable with a density $R(y) = \sup_{x|h(x)=y} Q(x)$.

The equivalent of the gaussian law is the *normal cost distribution*, a quadratic form $\mathcal{N}_{m,\sigma}(x) = -(1/2)(x-m)^2/\sigma^2$, or its vector form if the decision variable is a vector.

The equivalent of the convolution of probability measures is the sup-convolution, and the equivalent of the Fourier transform is the Fenchel transform. Using these tools, Quadrat was able to show a "law of large numbers" and a "central limit theorem". In turn, these induce original results on the limit behaviour of the return function of our dynamic programming example of paragraph 2.2.2, at least in the case where $L(x, y)$ is a function of $y - x$ only. This goes as follows (see [9]) :

Let a scalar simple system be given by

$$x_{t+1} = x_t + u_t ,$$

and a cost function to be maximized be given by

$$J(x_0, \{u_t\}) = K(x_N) + \sum_{t=0}^{N-1} L(u_t) .$$

Assume that L and K are concave C^2 functions, with 0 as their maximum. Let m be the point where L reaches its maximum : $L(m) = 0$, and let $M = (d^2L/du^2)(m) < 0$. Let $V_N(\cdot)$ be the return function of the N step maximization problem. Quadrat's central limit theorem states that

$$\lim_{N\to\infty} V_N\left(\sqrt{N}(x + Nm)\right) = -\frac{1}{2}Mx^2 .$$

3.2.2 Mathematical fear

Let φ be a decision variable over Ω. The natural equivalent to the mathematical expectation $\mathbb{E}\varphi$ of φ is the *mathematical fear*[3], denoted $\mathbb{F}\varphi$:

$$\mathbb{E}\varphi = \mathbb{E}^P_\omega \varphi := \int_\Omega \varphi(\omega) P(\omega)\, d\omega \,, \qquad \mathbb{F}\varphi = \mathbb{F}^Q_\omega \varphi := \sup_{\omega\in\Omega} [\varphi(\omega) + Q(\omega)] \,.$$

Of course, if φ depends on ω through another decision variable, say x of cost density Q_X, then we also have

$$\mathbb{F}\varphi = \mathbb{F}^Q_x \varphi = \sup_x [\varphi(x) + Q_X(x)] \,.$$

(It is a useful exercise at this stage to check that last formula, using the definition (1) of the cost density of the variable x, as this reasoning occurs very often in the use of cost measures.)

The first property to stress for the mathematical fear operator is that it is linear, in the $(\max, +)$ algebra of course. As a matter of fact, we do have that for two functions φ and ψ and a real number λ

$$\begin{aligned} \mathbb{F}(\max\{\varphi,\psi\}) &= \max\{\mathbb{F}\varphi, \mathbb{F}\psi\} \,, \\ \mathbb{F}(\lambda + \varphi) &= \lambda + \mathbb{F}\varphi \,. \end{aligned}$$

This will be of prominent importance in the sequel.

The second property we want to stress is that,

$$\text{if} \quad \forall \omega \in \Omega \,, \quad \varphi(\omega) \le \psi(\omega) \,, \qquad \text{then} \qquad \mathbb{F}\varphi \le \mathbb{F}\psi \,. \tag{2}$$

While this is clearly true from the definition of $\mathbb{F}$, it may be noticed that formally the reason is not the same as that of the same inequality for the expectation. The latter follows from the fact that the expectation of a nonnegative function is nonnegative and from the classical linearity of the expectation. But a substraction is involved in the proof, an operation with no equivalent in the new algebra. Here, the basic fact is that, if $\varphi(\omega) \le \psi(\omega)$, then $\max\{\varphi,\psi\} = \psi$, so that with $(\max, +)$ linearity,

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

hence the desired result. (No substraction is involved.)

[3]Quadrat et al. call it the mean, often denoted $\mathbb{M}$

4 Minimax optimization

4.1 Context

4.1.1 An alternative paradigm ?

Although this is not the feature we shall stress here, one should strongly object to the often heard assertion that "Nature is stochastic", in so far as the word "stochastic" refers to the use of probability theory. Nature is full of unpredictable events. Using the tools of probability to make a metaphor of them is *our choice*, not Nature's. Stated in other words, the famous probability space $(\Omega, \mathcal{A}, P)$ is in our heads, not in Nature. Constructing a cost space $(\Omega, \mathcal{B}, Q)$ (with $\mathcal{B}$ for the Borelian of a topology) is not more or less artificial. And, depending on the problem at hand and the modelization context, trying to minimize the fear of a performance index may be no less meaningful than trying to minimize an expectation.

Granted, the stochastic approach has a strong rooting in repeated experiments and average frequencies of occurences. This is directly related to the law of large numbers which *does* have an equivalent in cost measure theory. It remains to better understand if some situations, and which, take better advantage of this form of the law of large numbers, imposing an analysis in terms of cost measure, and a decision in mathematical fear. Some sort of Von Neumann's rationality axioms... There has been work in that direction, and recent work in artificial intelligence (e.g. [7, 11]), orin fuzzy measure theory (e.g. [13]) has revived earlier investigations of "games against nature" (see [8]) or of the foundations of statistics (see [12]).

4.2 A powerful tool

Independently of the motivation, we stress here that the concepts introduced so far constitute a powerful tool for the analysis of minimax decision problems. Indeed, a problem of the form

$$\min_u \mathbb{F} J(u, \omega)\,,$$

is in itself a minimax problem, since it is equivalent to

$$\min_u \sup_\omega [J(u, \omega) + Q(\omega)]\,.$$

As a consequence, the parallel with known approaches of stochastic control will be a way to derive results in minimax control. Moreover, the fundamental tool we want to extend to mathematical fear is that of dynamic programming, which lends itself to such generalizations as having state dependant cost densities and the like.

As a matter of fact, the derivation of a theory of minimax control in the absence of the certainty equivalence theorem (see e.g. [10]) was directly inspired by that parallel. But the first result formally obtained that way, as far as we know, is the one we recall below on L^1/L^∞ control [6], first reported in part at Sils Maria in 1997.

4.3 Dynamic minimax decision problems

4.3.1 The system

Let a discrete time partially observed disturbed control system be given by

$$x_{t+1} = f_t(x_t, u_t, w_t)\,, \tag{3}$$
$$y_t = h_t(x_t, w_t)\,, \tag{4}$$

where $x_t \in \mathbb{R}^n$ is the state vector at time t, u_t is the control vector at time t, to be chosen within a set $\mathsf{U} \subset \mathbb{R}^m$, $w_t \in \mathbb{R}^\ell$ is a disturbance vector at time t, may be constrained to belong to a set W, and $y_t \in \mathsf{Y} \subset \mathbb{R}^p$ is the observed output at time t.

We shall write $\mathbf{u} \in \mathcal{U}$ for the time sequence $\{u_t\}_{t\in[0,T-1]} \in \mathsf{U}^T$ (The upper index T is indeed a cartesian power, as it should, and *contrary* to the notations we introduce next and use in the rest of the paper) and similarly for $\mathbf{w} \in \mathcal{W}$ and $\mathbf{y} \in \mathcal{Y}$.

We shall need partial sequences defined as follows:

$$u^t = (u_0, u_1, \ldots, u_t)\,,$$

and similarly for all time sequences. (as a consequence, $\mathbf{u} = u^{[T-1]}$.) We shall let $u^t \in \mathsf{U}^t$ ([4]), $w^t \in \mathsf{W}^t$, $y^t \in \mathsf{Y}^t$.

Let also $\omega = (x_0, \mathbf{w})$ denote the disturbances a priori unknown to the controller, and $\omega \in \Omega = \mathbb{R}^n \times \mathcal{W}$. We also use $\omega^t = (x_0, w^t) \in \Omega^t = \mathbb{R}^n \times \mathsf{W}^t$.

We shall need the input-state and input-output maps of system (3)(4), that we call φ and η respectively, meaning that

$$x_t = \varphi_t(x_0, u^{t-1}, w^{t-1}) = \varphi_t(u^{t-1}, \omega^{t-1})\,, \tag{5}$$
$$y_t = \eta_t(x_0, u^{t-1}, w^{t-1}) = \eta_t(u^{t-1}, \omega^{t-1})\,. \tag{6}$$

[4] It is here that our notations are inconsistent, since U^t therefore stands for the cartesian power $t+1$ of U.

Finally, we shall use φ^t and η^t to mean the sequences $\{\varphi_\tau\}_{\tau=1,\dots,t}$ and $\{\eta_\tau\}_{\tau=1,\dots,t}$.

The problem shall always be to choose a control sequence to achieve a certain goal, based upon the knowledge of the noise corrupted output. And of course, the controller shall have to be causal, but with perfect recall: no past information is forgotten at any time. We shall even restrict it to be strictly causal. Thus an admissible strategy will be a sequence of maps $\{\mu_t : \mathsf{U}^{t-1} \times \mathsf{Y}^{t-1} \to \mathsf{U}\}_{t\in[0,T-1]}$ defining the control sequence through

$$u_t = \mu_t(u^{t-1}, y^{t-1})\,.$$

We shall let $\mathcal{M}$ denote the class of such admissible strategies.

To any admissible strategy and any $\omega \in \Omega$ corresponds a unique trajectory $\mathbf{x}$ and a unique control sequence $\mathbf{u}$. So that, although this is an abuse of notations, we shall write such things as $\varphi_T(\mu, \omega)$ where what we mean is the final state on the trajectory generated by that μ and ω.

4.3.2 The performance index

The set Ω is assumed to be endowed with a cost measure governing the decision variable ω. We assume that x_0 and $\mathbf{w}$ are independent, and that $\mathbf{w}$ is a white sequence, so that the cost measure is entirely specified by a cost density Q_0 over $\mathbb{R}^n$ governing x_0, and a sequence of cost densities $\{\Gamma_t\}$ over W governing the w_t's.

Remember also that cost densities are always normalized with their maximum at zero. We shall assume that all functions we use are upper semi continuous, and that the maxima are well defined. (For instance, the cost densities might have a compact domain.)

The natural equivalent to the classical performance index is

$$J(\mathbf{u}, \omega) = \max\{M(x_T)\,,\ \max_{0\le t<T} L_t(x_t, u_t, w_t)\} = \max_{0\le t\le T} L_t(x_t, u_t, w_t)\,, \qquad (7)$$

where we have, for convenience, let $L_T = M$, as we shall from now on.

Therefore, the criterion we shall strive to minimize will be

$$H(\mu) = \mathbb{F}J(\mu, \omega) \qquad (8)$$

It is worthwhile, to point out the following fact. We are interested in

$$\mathbb{F}J(u, \omega) = \mathbb{F}_{x_0}\mathbb{F}_{\mathbf{w}}J(\mathbf{u}, \omega) = \max_{x_0}\ \max_{w_0\dots w_{T-1}} \left[J(\mathbf{u}, \omega) + \sum_{k=0}^{T-1}\Gamma_k(w_k) + Q_0(x_0)\right].$$

The above expression involves the quantity $\mathbb{F}_{\mathbf{w}} J$ which can be expanded into

$$\mathbb{F}_{\mathbf{w}} J = \max_{w_0 \ldots w_{T-1}} \max_t \left[L_t(x_t, u_t, w_t) + \sum_{k=0}^{T-1} \Gamma_k(w_k) \right].$$

Now, this is equal to the same expression where we limit the summation sign to t instead of $T-1$:

Proposition 1

$$\mathbb{F}_{\mathbf{w}} J = \max_{w_0 \ldots w_{T-1}} \max_t \left[L_t(x_t, u_t, w_t) + \sum_{k=0}^{t} \Gamma_k(w_k) \right].$$

We leave the detailed verification of this proposition to the reader. It follows from the fact that after the maximizing t, w_t can always be chosen so as to maximize Γ, making it zero.

It is also useful to notice that this may be written in terms of

$$\bar{L}_t(x,u) := \mathbb{F}_w L_t(x,u,w) = \max_w [L_t(x,u,w) + \Gamma_t(w)].$$

We also have

Proposition 2

$$\mathbb{F}_{\mathbf{w}} J = \max_{\mathbf{w}} \max_t \left[\bar{L}_t(x_t, u_t) + \sum_{k=0}^{t-1} \Gamma_k(w_k) \right]. \tag{9}$$

This last form is usefull in that it shows that there is indeed no gain in generality in taking L_t to depend on w_t. We might as well consider only the problem in $\bar{L}^t$.

Finally, it only takes a carefull reading to check that in all the sequel, the Γ_t's may depend as well on x_t and u_t, without invalidating our calculations. So that although we shall write $\Gamma_t(w)$, the problem we consider is really to minimize over $\mathcal{M}$

$$H(\mu) = \mathbb{F}J(\mu,\omega) = \max_\omega \left\{ \max_{\tau \in [0,T]} L_\tau(x_\tau, u_\tau, w_\tau) + \sum_{\tau=0}^{T-1} \Gamma_\tau(x_\tau, u_\tau, w_\tau) + Q(x_0) \right\} \tag{10}$$

or any of the equivalent forms given by the propositions above. However, at this time, the Γ_t's are restricted to be normalized, i.e. $\max_w \Gamma_t(x,u,w) = 0$ for all (x,u). The reference [6] shows how to waive that restriction.

4.3.3 Perfect information

Let us first consider the simpler problem where the controller (choosing u) has access to the exact state, and therefore may control in state feedback. We can assume that x_0 is fixed, since it is known by the controller. It may always be made into a decision variable afterward.

We have an (extended) Isaacs equation:

$$\forall x \in \mathbb{R}^n\,, \quad V_T(x) = M(x)\,, \tag{11}$$

$$\forall t \in [0, T-1], \forall x \in \mathbb{R}^n\,, \quad V_t(x) = \inf_u \mathbb{F}_w^{\Gamma_t} \max\{V_{t+1}(f_t(x,u,w)), L_t(x,u,w)\} \tag{12}$$

We may state the following theorem

Theorem 1 *If the backwards recursion (11),(12) generates a bounded Value function V, then, the infimum of the problem (8) is given by $\mathbb{F}V_0(x_0)$ (recall that the initial state cost density Q_0 is given). Moreover, if the minimum in u is reached at $\varphi^*(t,x)$ in (12), then this is an optimal state feedback strategy.*

Proof Let us sketch the proof of the theorem. Let $\mathbf{u}$ be a fixed control sequence, and $\{x_t\}$, $t = 0, \ldots T$ the associated trajectory for a $\mathbf{w}$. We have

$$\begin{aligned}
V_{T-1}(x_{T-1}) &\le \mathbb{F}_w^{\Gamma_{T-1}} \max\{M(x_T), L_{T-1}(x_{T-1}, u_{T-1}, w)\} \\
V_{T-2}(x_{T-2}) &\le \mathbb{F}_w^{\Gamma_{T-2}} \max\{V_{T-1}(x_{T-1}), L_{T-2}(x_{T-2}, u_{T-2}, w)\} \\
\vdots &\le \vdots \\
V_1(x_1) &\le \mathbb{F}_w^{\Gamma_1} \max\{V_2(x_2), L_1(x_1, u_1, w)\} \\
V_0(x_0) &\le \mathbb{F}_w^{\Gamma_0} \max\{V_1(x_1), L_0(x_0, u_0, w)\}
\end{aligned}$$

There are two operations to perform on this sequence of inequalities. The first one is to take the fear of both sides of each with respect to the whole $\mathbf{w}$. This is possible preserving the inequalities thanks to (2). The fear appearing in the right hand side is in fact conditioned by the value taken by x_t. Even so, it is true that

$$\mathbb{F}_{\mathbf{w}} \mathbb{F}_w^{\Gamma_t} = \mathbb{F}_{\mathbf{w}}\,.$$

We shall come back on a deeper form of that property further. Now, taking it in this simple form, and using the linearity of the fear, it comes, in short hand notations (here $\mathbb{F}$ means $\mathbb{F}_{\mathbf{w}}$)

$$\begin{aligned}
\mathbb{F}V_{T-1} &\le \max\{\mathbb{F}M(x_T), \mathbb{F}L_{T-1}\} \\
\mathbb{F}V_{T-2} &\le \max\{\mathbb{F}V_{T-1}, \mathbb{F}L_{T-2}\} \\
\vdots &\le \vdots \\
\mathbb{F}V_1 &\le \max\{\mathbb{F}V_2, \mathbb{F}L_1\} \\
V_0 &\le \max\{\mathbb{F}V_1, \mathbb{F}L_0\}
\end{aligned}$$

The second operation is to get rid of the intermediate $\mathbb{F}V_t$'s. In the stochastic control case, one sums (equivalent to taking the max of the left sides and the max of the right hand sides above) and then substracts the terms appearing on both sides. We cannot do that here because we have no substraction. But instead, we may *substitute*. Substituting the second last inequality in the last one yields

$$V_0 \le \max\{\mathbb{F}V_2, \mathbb{F}L_1, \mathbb{F}L_0\},$$

and so on, until we end up with

$$V_0(x_0) \le \max_{t \in 1,\dots,T}\{\mathbb{F}L_t\},$$

with $L_T(x, u, w) = M(x)$ using (11). Use the proposition to conclude that a fortiori

$$V_0(x_0) \le \mathbb{F}_{\mathbf{w}} J(x_0, \mathbf{u}, \mathbf{w}). \tag{13}$$

But if u_t is chosen minimizing the r.h.s of (12), the $\le$ signs above are all replaced by $=$ signs, showing that that strategy yields $V_0(x_0) = J(x_0, \mathbf{u}, \mathbf{w})$ for the sequence $\mathbf{w}$ that provides the max at each step of the above procedure.

There remains to assume that u keeps using that state feedback strategy and chosing an arbitrary sequence $\mathbf{w}$ to have the opposite inequality signs in the above calculations, that reduce to equal signs if w choses the maximizing one, to conclude that indeed the value of this game is

$$V_0(x_0) = \mathbb{F}_{\mathbf{w}} J(x_0, \varphi^*, \mathbf{w}),$$

which, together with (13), concludes the proof upon taking the mathematical fear with respect to x_0 of both sides.

4.3.4 Imperfect information

We shall not develop the complete theory of the imperfect information case here, too long a topic for an introduction. The idea again is to follow the stochastic case according to the following program.

1. Compute in forward time a conditional state cost distribution $Q_t(x)$ using the available information.
2. Take that quantity and its dynamic equation (a "filter") as the new "state". It is known to the controller. Thus one can in principle develop a dynamic programming argument in the infinite dimensional space of state cost densities. Call $U_t(Q)$ the return function.
3. Check under what conditions the quantity $U_t(Q) := \mathbb{F}^Q V_t$ satisfies the dynamic programming equation. When it does, this yields a separation theorem.

The critical argument to carry out this program is as follows. As time goes on, the available information confines ω to belonging to a sequence of subsets of the perturbations *compatible with the available information* at each instant of time. In our case, the available information at time t is (u^{t-1}, y^{t-1}), and those *conditioning subsets* can be described as follows. Let

$$\Omega_t(x \mid u^{t-1}, y^{t-1}) = \left\{\omega \mid \varphi_t(u^{t-1}, \omega^{t-1}) = x \quad \text{and} \quad \eta^{t-1}(u^{t-1}, \omega^{t-1}) = y^{t-1}\right\}$$

The conditional state cost density of x at time t given the information u^{t-1}, y^{t-1} is the cost measure of that set, normalized by its maximum over $\mathbb{R}^n$. And the conditioning subset at time t is

$$\Omega_t = \Omega_t[u^{t-1}, y^{t-1}] := \bigcup_{x \in \mathbb{R}^n} \Omega_t(x \mid u^{t-1}, y^{t-1})$$

Whether the available information is generated through an instantaneous output as here or otherwise, the important feature is that it be full memory : nothing is forgotten, so that, *the sequence of conditioning sets Ω_t generated is decreasing*. (The equivalent of the imbedded algebras property of stochastic control.) If so, it is true that conditional fears compose as

$$\tau > t \quad \Longrightarrow \quad \mathbb{F}^{\Omega_t} \mathbb{F}^{\Omega_\tau} = \mathbb{F}^{\Omega_t} \, .$$

Thanks to that fact, we can indeed carry out the above program and end up with a valid theory. (We also need to ascertain that the available information is strictly *causal*, i.e. the condition $\omega \in \Omega_t$ only places constraints on the restriction ω^{t-1} of ω to $[1, \ldots, t-1]$.)

The full development of that theory, and its continuous time counterpart, as it stands now, can be found in [6].

5 Conclusion

At this stage, mathematical fear is above all a useful tool to extend to minimax control results and intuitions coming from the more ancient world of stochastic control, or of probability theory. Whether it will turn out to be more than that is a matter of conjecture, and of hard work !

Mathematical rigour has been dealt with in a light mood here. Although it is a concern, of course, things are by far less technical than in probability theory. As a matter of fact, the discrete time case raises little difficult questions. The situation is very different with the continuous time case. In [6], we assume that the Value function is smooth, an hypothesis known to be too restrictive. Since we really need the approach of dynamic programming followed here, we need more powerful tools of set valued analysis.

References

[1] M. Akian : "Densities of idempotent measures and large deviations", Transactions of the AMS **351** pp 4515–4543, 1999.

[2] F. Baccelli, G. Cohen, G-J. Olsder, and J-P. Quadrat : *Synchronization and linearity : an algebra for discrete event systems*, John Wiley & sons, New-York, 1992

[3] T. Başar and P. Bernhard : $\mathcal{H}_\infty$ *optimal control and related minimax design problems, a differential games approach*, Birkhauser, 1991 (second ed. 1995).

[4] E.N. Barron, P. Cardialaguet, and R. Jensen : "Radon-Nikodym theorems in L^∞", *Cahiers du Centre de Recherche Viabilité Jeux Contrôle*, Université Paris Dauphine, to appear.

[5] P. Bernhard : "A separation theorem for expected value and feared value discrete time control", *COCV*, **1**, pp 191–206, 1996.

[6] P. Bernhard: “Minimax —or feared value— L^1/L^∞ control”, to appear, *Theoretical Computer Science*, special issue on “Max-Plus Algebras”, 2000. (I3S internal report 1999)

[7] D. Dubois and H. Prade: “Possibility theory as a basis for qualitative decision theory”, *IJCAI'95*, pp 1925–1930, Montréal, Canada, 1995.

[8] R. Luce and H. Raiffa: *Games and decisions*, John Wiley & sons, New york, USA, 1957.

[9] J-P. Quadrat : “Théorèmes asymptotiques en programmation dynamique”, *CRAS*, **311**, pp 745–748, Paris, 1990.

[10] A. Rapaport et P. Bernhard: “Étude d'un jeu de poursuite plane avec connaissance imparfaite d'une coordonnée”, *APII* **29**, pp 575–602, 1995.

[11] R. Sabbadin : “Une approche ordinale de la décision dans l'incertain : axiomatisation, représentation logique et application à la décision séquentielle”, thèse, Université Paul Sabatier, Toulouse, France, 1998.

[12] L. J. Savage: *The foundations of statistics*, John Wiley & sons, New York, USA, 1954.

[13] M. Sugeno : “Fuzzy measures and fuzzy integrals”, in N.M. Gupta et al. ed.: *Fuzzy automata and decision processes*, pp 928–933, North Holland, Amsterdam, 1977.