Hahn-Banach Separation Theorem for Max-Plus Semimodules

Guy Cohen[*‡], Stéphane Gaubert[†‡], Jean-Pierre Quadrat[‡]

[*] *CERMICS, ENPC, 6–8, avenue Blaise Pascal, 77455 Marne-La-Vallée cedex 2, France*
[†] *UMA, ENSTA, 32 Boulevard Victor, 75739 Paris cedex 15, France*
[‡] *INRIA, Rocquencourt, B.P. 105, 78153 Le Chesnay cedex, France*

Abstract

We introduce max-plus analogues of basic Euclidian geometry notions: scalar product is replaced by a scalar division, and the associated distance is essentially Hilbert's projective distance. We introduce an orthogonal projection and prove a Hahn-Banach type theorem: a point can be separated from a semimodule by a hyperplane orthogonal to the direction of projection. We use these results to separate max-plus convex sets, and illustrate this new geometry by two-dimensional examples.

1 Introduction

In the last twenty years, the max-plus algebra and related structures (called "dioids" or idempotent semirings) have emerged as a natural setting for various areas of applied mathematics and mathematical models (e.g. the so-called "discrete event systems" area [1], optimization theory [13], etc.). As vector spaces are built up from fields of scalars, like $\mathbb{C}$ or $\mathbb{R}$, semimodules with an idempotent addition can be built up from the max-plus semiring. Such algebraic structures share several common features with their more conventional counterparts. The main departure point is the idempotency of addition which induces a semilattice structure, and often a lattice structure. This natural order provides an alternative way to solve problems which are more usually solved by appealing to the minus sign or division in conventional algebra. Residuation theory [2], the aim of which is to provide solution concepts to equations, is perhaps the best illustration of how order can help to provide for the absence of invertibility of operations such as addition and multiplication.

Since discrete event systems entered the realm of system theory in particular through the use of such algebraic tools, this is a good instance for making the following considerations. In linear system theory, most techniques (in particular for control synthesis purposes, but already for system description) have two origins: algebra and geometry (of linear vector spaces). This duality is well illustrated by the parallel works of Wolovich [15] and Wonham [16]. The confrontation of both points of view has been very fruitful for this area. Thanks to the algebraic tools alluded to before, similar (indeed very similar) developments have been made possible for discrete event systems, at least for the subcategory recognized as "linear systems" (which can also be viewed as timed event graphs in the Petri net parlance), and as far as *algebraic* techniques are concerned. But, admittedly, the *geometric* developments did not follow the same path and, in fact, our understanding of geometry in idempotent semimodules is more limited at this moment.

In the last few years, the authors made some progress in understanding some elementary geometric notions such as that of projection on the image of an operator parallel to the kernel of another operator [4]. These operators need not be linear — residuated is enough — and indeed residuation was the basic technique to provide expressions for projectors. Consequently, even when one starts from linear operators, projectors appear to be nonlinear if one wants to stick to either the max-plus or the min-plus algebra (both algebras need in fact be mixed in general). However, there are interesting cases when linearity of projectors can be preserved [5].

This paper is a continuation of this effort to understand some geometric notions in max-plus semimodules. As already said, projectors have been introduced so far using residuation techniques which are specific to lattice structures. In conventional Hilbert spaces and in convexity theory, *orthogonal* projections are defined by using such notions as scalar or duality product and minimization of norms. A related and very important topic is that of the Hahn-Banach theorem in its geometric form, namely the separation of nonintersecting convex sets. This result is the ground for many fundamental results in convexity ("external" representations of convex as intersections of "supporting" half spaces, subdifferential calculus for convex functions) and in optimization (duality and multipliers).

In the present paper, we will focus on what we believe to be analogues of these notions in max-plus semimodules. After a summary on the basic tools provided by residuation and on results obtained so far on projections, we will consider *orthogonal* projections, at least special projections which seem to play that role. Then, we will turn ourselves towards putting these notions of orthogonal projections in relation with some kind of "scalar products" allowing us to define kinds of "hyperplanes" which will then be used to state a "separation theorem". Despite the fact that our "scalar products" are not really products and that several other analogies we make may seem odd at first sight, we hope that these preliminary results will open the road to more progress in understanding the geometry of "subspaces" and analogues to "convex sets" in the framework of idempotent semimodules.

2 Summary about Residuation and Nonorthogonal Projections

2.1 Max-Plus Algebra and Idempotent Semimodules

The max-plus semiring, $\mathbb{R}_{\max}$, is the set $\mathbb{R} \cup \{-\infty\}$ equipped with the max operation as addition (denoted $\oplus$) and the conventional $+$ as multiplication (denoted $\otimes$, but this symbol is often omitted). The zero is $-\infty$, which is denoted ε. The usual 0, denoted e, is the unit for $\otimes$, and $\otimes$ distributes over $\oplus$. Finally, ε is absorbing for $\otimes$ ($\varepsilon \otimes x = \varepsilon$ for all x), and $\oplus$ is idempotent ($x \oplus x = x$). A natural order is associated with any idempotent operation, namely, $x \le y \Leftrightarrow y = y \oplus x$, and this order (here the usual order in $\mathbb{R} \cup \{-\infty\}$) is compatible with $\otimes$ (that is, all elements behave as "nonnegative" elements when multiplying both sides of inequalities). This natural order endows an idempotent semiring with a sup-semilattice structure (for which $x \oplus y = x \vee y$ is the least upper bound of x and y), and, in the case of $\mathbb{R}_{\max}$, it suffices to add $+\infty$ (denoted $\top$) to the set to obtain a *complete* sup-semilattice (in which arbitrary subsets have a least upper bound). The corresponding semiring will be denoted $\overline{\mathbb{R}}_{\max}$. It is a standard result that complete sup-semilattices are also (complete) lattices, which means that arbitrary subsets have a greatest lower bound (in particular, we denote $x \wedge y$ the greatest lower bound of $\{x, y\}$). In the case of $\overline{\mathbb{R}}_{\max}$, $\wedge$ is nothing but min. We say that an idempotent semiring is complete when it is complete as an ordered set, and when the product distributes over arbitrary sups. For instance, the semiring $\overline{\mathbb{R}}_{\max}$ is complete. (Notice that, in $\overline{\mathbb{R}}_{\max}$, since zero is absorbing, $\varepsilon \otimes \top = -\infty + \infty = \varepsilon = -\infty$.) It is straightforward to extend "addition" and "multiplication" to rectangular matrices. In particular, making the semiring of scalars $\overline{\mathbb{R}}_{\max}$ act on the additive monoid $(\overline{\mathbb{R}}^n_{\max}, \oplus)$ of n-dimensional columns vectors by multiplication, we equip $\overline{\mathbb{R}}^n_{\max}$ with a structure of (free, finitely generated) semimodule, in which addition is idempotent. We warn the reader that unlike vector spaces, idempotent semimodules are *not free*, except in very special cases. Although some of our results do hold for rather general semimodules, we shall only consider, in the sequel, finitely generated subsemimodules of the free semimodule $\overline{\mathbb{R}}^n_{\max}$: the main interesting features of the theory are already apparent in this case.

2.2 Residuation Theory

A mapping $f : \mathcal{U} \to \mathcal{X}$ between two ordered sets is *residuated* if it is isotone (that is, order-preserving), and if, for all $x \in \mathcal{X}$, the subset $\{u \in \mathcal{U} \mid f(u) \leq x\}$ admits a maximal element, denoted $f^\sharp(x)$. The isotone mapping $f^\sharp : \mathcal{X} \to \mathcal{U}$ is called the *residual* of f. The residual $f^\sharp$ is the only isotone mapping satisfying the following properties:[1]

$$f \circ f^\sharp \leq I \;, \qquad f^\sharp \circ f \geq I \;. \tag{1}$$

A simple characterization holds in the case of *complete* lattices. Before considering it, let us introduce some terminology.

When $\mathcal{U}$ and $\mathcal{X}$ are lattices, we say that $f : \mathcal{U} \to \mathcal{X}$ is a $\vee$- or sup-*morphism* if $f(u \vee v) = f(u) \vee f(v)$ for all $u, v \in \mathcal{U}$ (same terminology with $\wedge$). When the lattices $\mathcal{U}$ and $\mathcal{X}$ are complete, we say that f is $\vee$- or sup-*continuous* if f preserves least upper bounds of arbitrary sets (specializing this property to the empty set, we get $f(\varepsilon) = f(\sup \varnothing) = \sup \varnothing = \varepsilon$, where, ε denotes the bottom element of an ordered set). The dual property for $\wedge$ is called inf-continuity (in [1], these properties are called *lower* and *upper semicontinuity*, respectively). Finally, if $\mathcal{U}$ and $\mathcal{X}$ are semimodules, we say that f is *linear* if it is an additive morphism and, in addition, $f(\alpha u) = \alpha f(u)$ with α a scalar and $u \in \mathcal{U}$. Now, returning to our residuation summary, f is residuated iff f is sup-continuous. In particular, linear mappings between free finitely generated semimodules are residuated.

The following identities can be easily derived from (1):

$$f \circ f^\sharp \circ f = f \;, \quad f^\sharp \circ f \circ f^\sharp = f^\sharp \;, \quad (h \circ f)^\sharp = f^\sharp \circ h^\sharp \;, \tag{2}$$

where f, h are residuated mappings with $f : \mathcal{U} \to \mathcal{X}$, $h : \mathcal{X} \to \mathcal{Y}$.

The notion of *dually residuated* mapping is defined naturally by reversing the order in the above definitions. See [1] for details. We use the notation $f^\flat$ for the dual residual of f. An immediate consequence of characterization (1) and its dual is that a residuated map $f^\sharp$ is itself dually residuated and $\left(f^\sharp\right)^\flat = f$.

2.3 Matrix Residuation

In $\overline{\mathbb{R}}_{\max}$, consider the mapping $L_a : x \mapsto ax$ for some given a (L is for Left multiplication by a). This mapping is linear and thus residuated. Its residual $L_a^\sharp$ is denoted $y \mapsto a \backslash y$ (left "division" by a) and is actually the conventional subtraction of a from y with the additional rule (which results from the very definition): $\varepsilon \backslash \varepsilon = \top$ (that is, $-\infty + \infty = +\infty$, to be contrasted with $\varepsilon \otimes \top = \varepsilon$ which may also, ambiguously, be written as $-\infty + \infty = -\infty$).

[1]We denote I the identity map, without reference to the underlying set, which should be clear from the context.

Similar considerations apply to the left multiplication L_A by a rectangular matrix $A \in \overline{\mathbb{R}}_{\max}^{m\times n}$, with the following formula:

$$(A \backslash B)_{ik} = \bigwedge_{j=1}^{m} \left(A_{ji} \backslash B_{jk}\right) , \quad \text{for } 1 \le i \le n,\ 1 \le k \le p ,$$

where $B \in \overline{\mathbb{R}}_{\max}^{m\times p}$. Therefore, calculating $A \backslash B$ amounts to performing a kind of (left) matrix product of B by the *transpose* of A where scalar multiplication is replaced by (left) division and scalar addition is replaced by lower bound. Of course, since matrix product is not commutative, one must distinguish between *left* and *right* division, the latter, denoted $\cdot \,/\, A$, being the residual of right multiplication $R_A(\cdot) = \cdot \otimes A$. We shall use the following general residuation inequalities (see [1, Table 4.1]), which hold in particular for rectangular matrices (of compatible dimensions):

$$A(A \backslash B) \le B , \tag{3a}$$
$$(A \backslash B)C \le A \backslash (BC) . \tag{3b}$$

One must be careful in using expressions such as $A \backslash Bx$ which, as written without parentheses, are ambiguous. On the one hand, when for instance $x \in \overline{\mathbb{R}}_{\max}^{p}$, $A \backslash (Bx)$ is interpreted as $L_A^\sharp \circ L_B(x)$: $L_A^\sharp \circ L_B$ is not in general a $\oplus$-morphism from $\overline{\mathbb{R}}_{\max}^{p}$ to $\overline{\mathbb{R}}_{\max}^{n}$. On the other hand, $x \mapsto (A \backslash B)x$ is to be interpreted as a *linear* operator from $\overline{\mathbb{R}}_{\max}^{p}$ to $\overline{\mathbb{R}}_{\max}^{n}$ because $A \backslash B$ is, by definition, the greatest matrix X such that $AX \le B$. In terms of operators, one can prove that $L_A^\sharp \circ L_B \ge A \backslash B$ using (3b).

3 Projections

3.1 Nonorthogonal Projections

This section summarizes results published in [4, 5] on projections on $\operatorname{im} B$ parallel to the kernel of C (denoted $\ker C$), where $B : \mathcal{U} \to \mathcal{X}$ and $C : \mathcal{X} \to \mathcal{Y}$ are residuated or linear operators between complete semimodules (say, here, $\mathcal{U} = \overline{\mathbb{R}}_{\max}^{m}$, $\mathcal{X} = \overline{\mathbb{R}}_{\max}^{n}$, $Y = \overline{\mathbb{R}}_{\max}^{p}$). In semimodules, it makes sense to define $\ker C$ as the following *equivalence relation* over $\mathcal{X}$:

$$x \overset{\ker C}{\sim} \xi \Leftrightarrow C(x) = C(\xi) \Leftrightarrow x \in C^{-1}\big(C(\xi)\big) , \tag{4}$$

rather than in the more usual way $\{x \in \mathcal{X} \mid C(x) = \varepsilon\}$ which is not very useful. The *projection* ξ *of* $x \in \mathcal{X}$ *on* $\operatorname{im} B$ *parallel to* $\ker C$ is such that $\xi \in \operatorname{im} B$ and $\xi \overset{\ker C}{\sim} x$. Compared with the analogous notion in conventional vector spaces, one must consider that, in a way, $\xi - x \in \ker C$ is the direction of projection, but notice how the absence of a "minus sign" is now compensated for. As in the classical case, there are conditions for existence and uniqueness of such projections, and then one can possibly get an explicit formula for the corresponding projector in terms of B and C (in linear vector spaces and for matrices B and C such that CB is invertible, the projector Π_B^C is equal to $B(CB)^{-1}C$).

In the present situation, conditions for existence and uniqueness are also known (see [4]) under which Π_B^C is given by the expression

$$\Pi_B^C = B \circ (C \circ B)^\sharp \circ C = (B \circ B^\sharp) \circ (C^\sharp \circ C) . \tag{5}$$

Observe that the former expression has a strong similarity with that encountered in vector spaces, whereas the latter form is written as the composition of two projectors, one on $\operatorname{im} B$

and the other one parallel to $\ker C$. These individual projectors are *orthogonal* projectors which are discussed in the next section.

Even if either the uniqueness or the existence (or both) condition(s) is (are) not satisfied, the operator Π_B^C, defined by (5), acts in an interesting way: $\Pi_B^C(x)$ is the maximal element ξ of the image of B such that $C\xi \le Cx$ (but equality does hold true if x is already in $\operatorname{im} B$).

If B and C are not only residuated but *linear*, and if the uniqueness and existence conditions are both satisfied, then, it turns out that Π_B^C is indeed *linear*. More explicitly, the expressions in (5) which are seemingly nonlinear, boil down to either of the following two forms:

$$\Pi_B^C = \big(B \,/\, (CB)\big)C = B\big((CB) \backslash C\big) . \tag{6}$$

The question arises of when, for a given subsemimodule $\operatorname{im} B$ (B and C are still supposed linear here), there exists a subsemimodule of the form $\ker C$ which is such that the existence and uniqueness conditions are satisfied (then the projector is linear as just said), in which case we say that $\operatorname{im} B$ and $\ker C$ are *direct factors*.[2] We showed in [5] that this property holds iff B is *regular*, that is, if B has a g-inverse $B^\dagger$, which satisfies $BB^\dagger B = B$. Then, the maximal g-inverse is $B \backslash B / B$, and B has a g-inverse iff $B = B(B \backslash B / B)B$, which allows us to check regularity (the expression $B \backslash B / B$ is nonambiguous because, in general, $(U \backslash V) / W = U \backslash (V / W)$).

3.2 Orthogonal Projections

Let $B : \mathcal{U} \to \mathcal{X}$ be a residuated operator with $\mathcal{U} = \overline{\mathbb{R}}_{\max}^m$ and $\mathcal{X} = \overline{\mathbb{R}}_{\max}^n$. The following theorem provides equivalent definitions of the *orthogonal projection* Π_B on $\operatorname{im} B$.

Theorem 1. *Let* $\Pi_B \stackrel{\text{def}}{=} B \circ B^\sharp$. *Then,*

- $\xi = \Pi_B(x)$ *is the* greatest *element in* $\operatorname{im} B$ *which is* less *than* x.
- Π_B *is the projector on* $\operatorname{im} B \subset \mathcal{X}$ *parallel to* $\ker B^\sharp \subset \mathcal{X}$.

Proof. Looking for ξ such that $\xi = Bz$ for some z and $Bz \le x$, we know that the greatest solution is provided by $z = B^\sharp(x)$, hence $\xi = \Pi_B(x)$, which proves the former statement. Also, $B^\sharp(\xi) = B^\sharp(x)$ (from (2)) which shows that the projection is parallel to $\ker B^\sharp$. $\square$

Of course, if $x \in \operatorname{im} B$, $\Pi_B(x) = x$, which shows that $\operatorname{im} \Pi_B = \operatorname{im} B$ and $\Pi_B \circ B = B$.

We call Π_B an *orthogonal* projector because in standard algebra, with B a matrix, $\ker B^\top$ (the transpose of B) is orthogonal to $\operatorname{im} B$, and we believe that $B^\sharp$ plays the role of $B^\top$ in our context. This terminology will also be enforced by the results to come on the separation theorem.

4 Max-plus Inversion, Scalar Division, and Hilbert's Projective Metric

In $\overline{\mathbb{R}}_{\max}$, we consider the transformation: $x \mapsto x^- \stackrel{\text{def}}{=} e / x = x \backslash e$. Note that $(x^-)^- = x$, $e^- = e$ and $\varepsilon^- = \top$. When $x \in \overline{\mathbb{R}}_{\max}^{n\times p}$, we set $x^- = \Phi_p / x = x \backslash \Phi_n$, where for all $k \ge 1$, Φ_k denotes the $k \times k$ matrix:

$$\Phi_k \stackrel{\text{def}}{=} \begin{pmatrix} e & \top & \dots & \top \\ \top & e & \ddots & \vdots \\ \vdots & \ddots & \ddots & \top \\ \top & \dots & \top & e \end{pmatrix} .$$

[2] The role of the "given" and the "whether there exists" operators can be inverted, that is, the property is a symmetric one between the image and the kernel subsemimodules.

In the sequel, we shall simply write Φ instead of Φ_k, and we denote e the identity matrix of any size. The practical rule to compute x^- is: *transpose* the matrix and *inverse* its entries. Again $(x^-)^- = x$, and $e^- = \Phi$, $\varepsilon^- = \top$. We next list some useful properties of the inversion $x \mapsto x^-$, which follow from the general formulæ of [1, Table 4.1]. First $(a \oplus b)^- = a^- \wedge b^-$, hence $(a \wedge b)^- = a^- \oplus b^-$. In particular, the mapping $a \mapsto a^-$ is *antitone, residuated, dually residuated* — in an adapted sense — and *"self-residuated"* (both residuals are equal to the mapping itself). Also, $(ab)^- = b \backslash a^- = b^- \not/ a$, hence $b \backslash a = (a^- b)^-$ and $a \not/ b = (ba^-)^-$ (avoid division!). Moreover, $aa^- \leq \Phi$. For example,

$$a = \left(\begin{smallmatrix} e & \varepsilon \\ \varepsilon & \varepsilon \end{smallmatrix}\right), \quad a^- = \left(\begin{smallmatrix} e & \top \\ \top & \top \end{smallmatrix}\right), \quad aa^- = \left(\begin{smallmatrix} e & \top \\ \varepsilon & \varepsilon \end{smallmatrix}\right) .$$

However, $a \backslash \big(aa^-\big) = \big(a^- a\big) \not/ a = a^-$.

We now play with *column* vectors with entries in $\overline{\mathbb{R}}_{\max}$. We define $\langle x \mid y \rangle \overset{\text{def}}{=} x^- y$ and $[x \mid y] \overset{\text{def}}{=} x \backslash y$. We shall need the following easy properties, that we state without proof:

1. $\langle x \mid y \rangle = (y \backslash x)^- = (y^- \not/ x^-)^-$ and $[x \mid y] = \langle y \mid x \rangle^-$;
2. for all $x \in \overline{\mathbb{R}}^n_{\max}$, $\langle x \mid x \rangle \leq e$ and $[x \mid x] \geq e$;
3. for scalars α,

$$\langle x \mid y \rangle \, \alpha = \langle x \not/ \alpha \mid y \rangle = \langle x \mid y\alpha \rangle \qquad \text{and} \qquad \alpha \backslash [x \mid y] = [x\alpha \mid y] = \langle x \mid y \not/ \alpha \rangle \ ;$$

4. for all matrices A, $\langle x \mid Ay \rangle = \big\langle A^\sharp(x) \mid y \big\rangle$ and $\big[x \mid A^\sharp(y)\big] = [Ax \mid y]$;
5. the following three statements are equivalent:
   (a) $\langle x \mid x \rangle = e$,
   (b) $[x \mid x] = e$,
   (c) x has at least one finite coordinate ($\neq \varepsilon$ and $\top$).
6. If $\langle y \mid x \rangle \leq e$ and $y \leq x$, then $y = x$.

When x, y have *finite* entries, the scalar products $\langle \cdot \mid \cdot \rangle$ and $[\cdot \mid \cdot]$ have a remarkable geometric interpretation: $\langle x \mid y \rangle \oplus \langle y \mid x \rangle = \|x - y\|_\infty$, where $\|\cdot\|_\infty$ denotes the sup-norm, and $\langle x \mid y \rangle \otimes \langle y \mid x \rangle = \|x - y\|_H$, where $\|x\|_H = \max_{1 \leq i \leq n} x_i - \min_{1 \leq i \leq n} x_i$ denotes Hilbert's seminorm. Hilbert's seminorm induces a norm on the additive projective space, which is the quotient of $\mathbb{R}^n$ by the additive parallelism relation: $x \parallel y \iff x - y$ is a constant vector $\iff x = \lambda \otimes y$ for some $\lambda \in \mathbb{R}$. In the sequel, it will be convenient to extend the definition of the associated Hilbert's additive projective distance, $d_H(x, y) = \|x - y\|_H$, to the case when x, y have infinite values. The right definition turns out to be:

$$d_H(x, y) = \big([x \mid y] \otimes [y \mid x]\big)^-$$

(since for scalars a, b, $(ab)^- = a^- b^-$ except in the exceptional case when $(a, b) = (\varepsilon, \top)$ or $(\top, \varepsilon)$, we see in particular that $d_H(x, y)$ coincides with $\langle x \mid y \rangle \otimes \langle y \mid x \rangle$ when the entries of x and y are finite). Our generalization d_H of Hilbert's distance satisfies the following properties:

$$d_H(x, z) \leq \big(d_H(x, y)^- \otimes d_H(y, z)^-\big)^- \qquad \text{(triangular inequality)} \qquad (7a)$$
$$d_H(x, y) = 0 \implies x = \lambda \otimes y \text{ for some } \lambda \in \mathbb{R} \qquad \text{(definiteness)} \qquad (7b)$$
$$d_H(x, y) < 0 \iff x = y \in \{\varepsilon, \top\}^n \qquad \text{(nonnegativity)} \qquad (7c)$$

(the converse implication holds in (7b) when $x, y \notin \{\varepsilon, \top\}^n$).

5 Bivectors, Duality and Separation Theorem

We call *bivector* a pair of vectors (x, y). The *orthogonal* of the bivector $(x, y) \in (\overline{\mathbb{R}}^n_{\max})^2$ is the semimodule:

$$(x, y)^\perp \stackrel{\text{def}}{=} \left\{ z \in \overline{\mathbb{R}}^n_{\max} \;\middle|\; \langle x \mid z\rangle = \langle y \mid z\rangle \right\} .$$

Theorem 2. *Given a linear operator* $B : \overline{\mathbb{R}}^p_{\max} \to \overline{\mathbb{R}}^n_{\max}$ *and* $x \in \overline{\mathbb{R}}^n_{\max}$, $\Pi_B(x)$ *is the* least $\xi \in \overline{\mathbb{R}}^n_{\max}$ *such that*

$$\operatorname{im} B \subset (x, \xi)^\perp .$$

Proof. Since $\langle x \mid B(u)\rangle = \langle \xi \mid B(u)\rangle$ iff $\langle B^\sharp(x) \mid u\rangle = \langle B^\sharp(\xi) \mid u\rangle$, $\operatorname{im} B \subset (x, \xi)^\perp$ iff $B^\sharp(x) = B^\sharp(\xi)$. The least such ξ is $B\bigl(B^\sharp(x)\bigr)$ because $B^\sharp$ is dually residuated with $(B^\sharp)^\flat = B$. □

Theorem 3. *The bivector* $\bigl(x, \Pi_B(x)\bigr)$ *separates* $\operatorname{im} B$ *from* x *iff* $x \not\in \operatorname{im} B$.

Proof. We know from Theorem 2 that $\operatorname{im} B \subset \bigl(x, \Pi_B(x)\bigr)^\perp$. It remains to prove that x itself is *not* orthogonal to that bivector iff it does not belong to $\operatorname{im} B$, that is,

$$\langle x \mid x\rangle \neq \langle \Pi_B(x) \mid x\rangle \iff x \not\in \operatorname{im} B .$$

Indeed, it suffices to prove that $\langle x \mid x\rangle = \langle \Pi_B(x) \mid x\rangle \Rightarrow x \in \operatorname{im} B$ (which is equivalent to saying that $x = \Pi_B(x)$). If $\langle x \mid x\rangle = \langle \Pi_B(x) \mid x\rangle$, then $\langle \Pi_B(x) \mid x\rangle \le e$ according to item 2 of §4. Moreover, $\Pi_B(x) \le x$ (see Theorem 1), hence both assumptions in item 6 of §4 are satisfied and we get that $\Pi_B(x) = x$. □

As an immediate corollary of the separation theorem, we get the following duality result, a variant of which was already proved in [7, Ch. 3, Cor.1.2.5] (see also [9, Th. 9]). The *orthogonal* $\operatorname{im} B^\top$ is the set of bivectors $(y, z) \in (\overline{\mathbb{R}}^n_{\max})^2$ such that $\langle y \mid u\rangle = \langle z \mid u\rangle$ for all $u \in \operatorname{im} B$.

Corollary 4. *We have* $(\operatorname{im} B^\top)^\perp = \operatorname{im} B$.

Theorem 3 should be geometrically intuitive: in an Euclidian space, to separate a point x from a convex set $\mathcal{B}$, a canonical choice is to take an hyperplane orthogonal to the vector (x, x'), where x' is the projection of x onto $\mathcal{B}$. Moreover, the projection minimizes the Euclidian distance. We next give a max-plus analogue of the later property: the point $\Pi_B(x)$ which defines the direction of the separating hyperplane in Theorem 3 minimizes the generalized Hilbert's projective distance d_H.

Theorem 5. *For all* $x \in \overline{\mathbb{R}}^n_{\max}$ *and* $y \in \operatorname{im} B$, $d_H(x, y) \ge d_H(x, \Pi_B(x))$.

Proof. Setting $y = Bu$ and using the inequalities (3) together with items 2 and 4 of §4, we get

$$\begin{aligned}
d_H(x, Bu)^- &= [x \mid Bu][Bu \mid x] \\
&= [x \mid Bu]\bigl[u \mid B^\sharp(x)\bigr] = \bigl(x \backslash (Bu)\bigr)\bigl(u \backslash B^\sharp(x)\bigr) \\
&\le x \backslash \Bigl(Bu\bigl(u \backslash B^\sharp(x)\bigr)\Bigr) \le x \backslash (BB^\sharp(x)) \\
&\le \Bigl(x \backslash \bigl(BB^\sharp(x)\bigr)\Bigr)\bigl(B^\sharp(x) \backslash B^\sharp(x)\bigr) \\
&= \bigl[x \mid BB^\sharp(x)\bigr]\bigl[BB^\sharp(x) \mid x\bigr] = d_H(x, \Pi_B(x))^- \;,
\end{aligned}$$

and after inversion, we obtain $d_H(x, Bu) \ge d_H(x, \Pi_B(x))$. □

6 Max-plus Affine Spaces and Convex Sets

In this section, we illustrate the above constructions with max-plus convex sets, using the classical correspondence between projective and affine geometry, which sends convex cones to convex sets.

A max-plus convex combination of two vectors $u, v \in \overline{\mathbb{R}}_{\max}^{n}$ is any vector of the form $\alpha u \oplus \beta v$, where $\alpha \oplus \beta = e$. A subset $\mathcal{B} \subset \overline{\mathbb{R}}_{\max}^{n}$ is *convex* if it is stable by convex combinations. We say that $\mathcal{B}$ is *finitely generated* if $\mathcal{B} = \mathrm{vex} B$ for some $B \in \overline{\mathbb{R}}_{\max}^{n \times p}$, where $\mathrm{vex} B = \{Bu \mid \bigoplus_i u_i = e\}$ denotes the set of convex combinations of the columns of a matrix B. We denote by $\hat{B} \in \overline{\mathbb{R}}_{\max}^{(n+1)\times p}$ the matrix obtained by adding a row of unit elements to B. With any convex set $\mathcal{B} \subset \overline{\mathbb{R}}_{\max}^{n}$, we associate the semimodule $\hat{\mathcal{B}} \subset \overline{\mathbb{R}}_{\max}^{n+1}$ generated by the vectors of the form $\hat{u}$, where $u \in \mathcal{B}$. In particular, when $\mathcal{B} = \mathrm{vex} B$ is finitely generated (as a convex set), $\hat{\mathcal{B}} = \mathrm{im}\, \hat{B}$ is finitely generated (as a semimodule).

The results of the above section allow us to separate a point from a finitely generated convex set, by means of affine hyperplanes. Indeed, if x does not belong to the convex set $\mathrm{vex} B$, $\hat{x}$ does not belong to the semimodule $\mathrm{im}\, \hat{B}$, and by Theorem 3, the bivector $(\hat{x}, \Pi_{\hat{B}}(\hat{x}))$ separates $\hat{x}$ from $\mathrm{im}\, \hat{B}$. This result can be translated in affine terms by introducing *affine hyperplanes*, which are sets of points $z \in \overline{\mathbb{R}}_{\max}^{n}$ solutions of $\langle a \mid z \rangle \oplus \alpha = \langle b \mid z \rangle \oplus \beta$, for some $a, b \in \overline{\mathbb{R}}_{\max}^{n}$ and $\alpha, \beta \in \overline{\mathbb{R}}_{\max}$ such that $(a, \alpha) \neq (b, \beta)$. Since

$$\Pi_{\hat{B}}(\hat{x}) = \begin{pmatrix} B(B^{\sharp}(x) \wedge E) \\ E^t(B^{\sharp}(x) \wedge E) \end{pmatrix} ,$$

where E denotes the p dimensional column vector whose entries are equal to e, let us introduce the affine hyperplane

$$H(x) = \Big\{ z \in \overline{\mathbb{R}}_{\max}^{n} \mid \langle x \mid z \rangle \oplus e = \big\langle B(B^{\sharp}(x) \wedge E) \mid z \big\rangle \oplus \big\langle E^t(B^{\sharp}(x) \wedge E) \mid e \big\rangle \Big\} . \tag{8}$$

Theorem 3 implies that $H(x)$ separates x from $\mathrm{vex} B$, i.e $H(x) \supset \mathrm{vex} B$ but $x \notin H(x)$. Moreover, Theorem 1 shows that the operator Π'_B defined by

$$\Pi'_B(x) = \big(E^t(B^{\sharp}(x) \wedge E)\big)^{-1} B(B^{\sharp}(x) \wedge E) \tag{9}$$

for all $x \in \overline{\mathbb{R}}_{\max}^{n}$ such that $B^{\sharp}(x) \neq \varepsilon$, is a projector onto $\mathrm{vex} B$.

For instance, the convex set generated by the columns of the matrix $B = \left(\begin{smallmatrix} 0 & 1 & 3 \\ 0 & 3 & 4 \end{smallmatrix}\right)$ is the dark region depicted in Figure 1. The three columns are the extremal points P, N, M of the

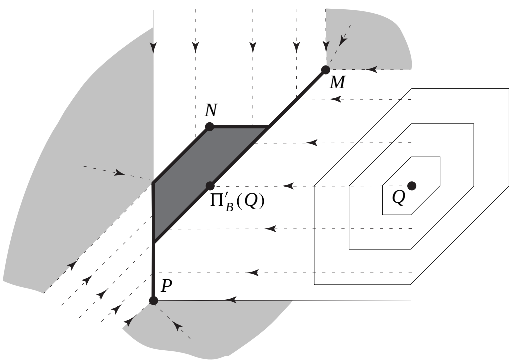

Figure 1: The convex generated by the 3 points (M,N,P) and the action of the projector.

convex set. The projector on $\hat{B}$ is

$$\Pi_{\hat{B}}((x,y,z)) = \begin{pmatrix} (x \wedge y \wedge z) \oplus (x \wedge (-1)y \wedge 3z) \\ (2x \wedge y \wedge 3z) \oplus (1x \wedge y \wedge 4z) \\ x \wedge y \wedge z \end{pmatrix}$$

and the action of the projector Π'_B on vex$B$ is represented by arrows. In this figure, we see that the point Q is sent to $\Pi'_B(Q)$ and that $(\Pi'_B)^{-1}(P)$ is the shaded region with vertex P. We have represented some balls of center Q for the distance d'_H obtained by transporting Hilbert's projective distance to the affine space: $d'_H(x,y) = d_H(\hat{x},\hat{y})$. Since $\Pi_{\hat{B}}(\hat{x})$ minimizes the Hilbert's projective distance from $\hat{x}$ to im $\hat{B}$, $\Pi'_B(x)$ minimizes the distance d'_H from x to vex$B$, a property that is geometrically clear from the shape of the balls.

Before considering separating hyperplanes, it is useful to look at the geometry of affine max-plus hyperplanes of $\overline{\mathbb{R}}^2_{\max}$, that we shall call *lines*. The general line is defined by an equation of the form $ax \oplus by \oplus c = a'x \oplus b'y \oplus c'$, for some $a, b, c, a', b', c' \in \overline{\mathbb{R}}_{\max}$, but not so many coefficients are needed. For instance, the lines with equations $2x \oplus y = 1x \oplus y \oplus 3$ and $2x \oplus y = y \oplus 3$, coincide. More generally, it is not difficult to see that there are 12 generic shapes of lines, as shown in Figure 2. Indeed, a generic line can be defined by three real

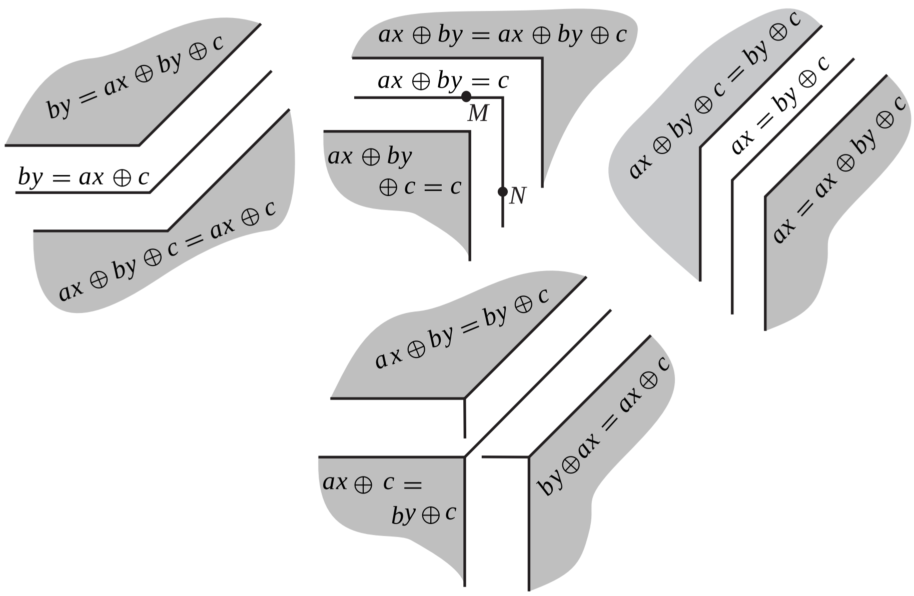

Figure 2: The twelve generic lines of $\overline{\mathbb{R}}^2_{\max}$

numbers a, b, c plus a "sign" information, which tells the side of the equation in which the corresponding coefficients is dominant (say "$\oplus$" for the left hand side, "$\ominus$" for the right and side, and a dot when coefficients on both sides are equal). For instance, the line with equation $ax \oplus c = by \oplus c$ will be denoted $L(\oplus a, \ominus b, \dot{c})$. This notation can be justified by introducing the *symmetrized* max-plus semiring [12, 1]. It is fundamental to note that a line with a dotted coefficient has dimension 2 in the usual sense. There is no point to distinguish algebraically between lines and half-planes, since for instance an inequality of the form $x \geq y$ can be written as an equation $x = x \oplus y$. Coming back to our example, the separating line $H(Q)$, given by (8), is $L(1, \dot{2}, \dot{2}) = \{(x,y) \mid 1x \oplus y \oplus 2 = y \oplus 2\}$.

Corollary 4 can be rephrased by saying that a (finitely generated) semimodule is exactly the set of solutions of the linear equations that it satisfies. Translating this theorem to the affine case, we get in particular that the convex vex$B$ is the intersection of the lines in which it is contained. In fact, it is not difficult to see that vex$B$ is the intersection of the five following lines:

$$L(\dot{1}, 0, \dot{3}) : 1x \oplus y \oplus 3 = 1x \oplus 3, \qquad L(0, -\dot{1}, \dot{0}) : x \oplus (-1)y \oplus 0 = (-1)y \oplus 0,$$
$$L(\dot{0}, -2, 0) : x \oplus (-2)y \oplus 0 = x, \quad L(0, \varepsilon, \dot{3}) : x \oplus 3 = 3, \quad L(\varepsilon, \dot{0}, 0) : y \oplus 0 = y \ .$$

The first line, $L(\dot{1}, 0, \dot{3})$ is the half-plane whose upper boundary contains the segment (M, N), the second line, $L(0, -\dot{1}, \dot{0})$, has a lower boundary which contains the segment (P, M), whereas the third line, $L(\dot{0}, -2, 0)$, has upper boundary (P, N). The two remaining lines make vertical and horizontal cuts at M and P, respectively.

More generally, a finitely generated convex set is the intersection of finitely many hyperplanes. Passing from the set of generators of a convex set to a definition as an intersection of hyperplanes is a non trivial operation: the only known algorithm is nonpolynomial [3], [7, Ch. 3] (see also [9, 8]).

Acknowledgement Upon completion of this paper, the authors learned the existence of the reference [11] which contains related results. We also learned that K. Zimmerman proved a separation theorem [17], but we could not obtain this reference yet. Finally, we thank P. Lotito and E. Mancinelli for useful discussions around the inversion introduced in section 4.

References

[1] F. Baccelli, G. Cohen, G.J. Olsder, J.-P. Quadrat. *Synchronization and Linearity — An Algebra for Discrete Event Systems*. Wiley, New-York, 1992.

[2] T.S. Blyth, M.F. Janowitz. *Residuation Theory*. Pergamon Press, 1972.

[3] P. Butkovič and G. Hegedüs. An elimination method for finding all solutions of the system of linear equations over an extremal algebra. *Ekonomicko-matematicky Obzor*, 20, 1984.

[4] G. Cohen, S. Gaubert, J.P Quadrat. Kernels, images and projections in dioids. *IEE Proceedings of WODES96*, Edinburgh, August 1996.

[5] G. Cohen, S. Gaubert, J.P Quadrat. Linear Projectors in the Max-Plus Algebra. *5th IEEE Mediterranean Conference on Control and Systems*, Paphos, Cyprus, 21–23 July, 1997.

[6] R.A. Cuninghame-Green. Minimax Algebra. *Lecture Notes in Econ. and Math. Systems N. 166* Springer-Verlag, 1979.

[7] S. Gaubert. Théorie des systèmes linéaires dans les dioïdes. *Thèse Ecole des Mines de Paris*, 1992.

[8] S. Gaubert, Two Lectures on Max-Plus Algebra, *in : Support de cours de la 26 ième École de Printemps d'Informatique Théorique,* Noirmoutier, May 1998.

[9] S. Gaubert and M. Plus. Methods and applications of (max,+) linear algebra. R. Reischuk and M. Morvan, editors, *STACS'97*, number 1200 in LNCS, Lübeck, March 1997. Springer.

[10] G. L. Litvinov, V. P. Maslov, and G. B. Shpiz. Linear Functionals on Idempotent Spaces: An Algebraic Approach. *Doklady Mathematics*, vol. 58, No. 3, 1998, pp. 389–391.

[11] G. L. Litvinov, V. P. Maslov, and G. B. Shpiz. Idempotent functional analysis: An algebraical approach. E-print math.FA/0009128, Sep. 2000, available from http://arXiv.org

[12] M. Plus. Linear systems in $(\max, +)$-algebra. *Proceedings of the 29th Conference on Decision and Control*, Honolulu, Dec. 1990.

[13] J.P. Quadrat, Max-Plus Working Group. Max-Plus Algebra and Applications to System Theory and Optimal Control. *Proceedings of the International Congress of Mathematicians,* Zurich, 1994, Birkhaüser , 1995.

[14] E. Wagneur. Moduloids and pseudomodules. 1. dimension theory. *Discrete Math.*, 98:57–73, 1991.

[15] W.A. Wolovich. *Linear Multivariable Systems*. Springer-Verlag, Berlin, 1974.

[16] W.M. Wonham. *Linear multivariable control: a geometric approach*. Springer-Verlag, Berlin, 2nd ed, 1979.

[17] K. Zimmermann. A general separation theorem in extremal algebras. *Ekonom.-Mat. Obzor*, 13(2):179–201, 1977.