Algèbre max-plus, applications monotones contractantes et équations aux dérivées partielles : trois approches du contrôle optimal

par

Marianne Akian

Mémoire pour l'obtention de

l'*Habilitation à Diriger des Recherches*

spécialité *Mathématiques*

délivrée par l'Université Pierre et Marie Curie (Paris 6)

Soumis aux *rapporteurs* :

M John MALLET-PARET	Professeur, Brown University
M Geert Jan OLSDER	Professeur, Delft University of Technology
M Marc QUINCAMPOIX	Professeur, Université de Brest

Présenté le 13 novembre 2007 devant le *jury* composé de :

M Geert Jan OLSDER	
M Marc QUINCAMPOIX	
Mme Nicole EL KAROUI	Professeur, École Polytechnique
M Albert FATHI	Professeur, École Normale Supérieure de Lyon
M Pierre-Louis LIONS (président)	Professeur, Collège de France
M Jean-Pierre QUADRAT	Directeur de Recherche, INRIA
M Alain ROUAULT	Professeur, Université de Versailles Saint-Quentin
M Sylvain SORIN	Professeur, Université Pierre et Marie Curie

Résumé

La plupart des travaux présentés ici sont reliés à l'étude de problèmes de contrôle optimal déterministe ou stochastique par la méthode de la programmation dynamique.

L'algèbre max-plus ou tropicale permet de voir les opérateurs de la programmation dynamique associés aux problèmes de contrôle optimal déterministe non actualisés comme des applications linéaires. Ceci nous a conduit à étudier certains systèmes linéaires et problèmes spectraux dans l'algèbre max-plus, à développer une théorie des probabilités idempotentes ainsi qu'une méthode des éléments finis max-plus.

Plus généralement, les opérateurs de la programmation dynamique sont des applications monotones, contractantes au sens large pour la norme L^∞. Nous avons obtenu plusieurs résultats en théorie de Perron-Frobenius non linéaire (problème spectral, asymptotique des itérées d'applications), et fait le lien avec l'étude de problèmes de contrôle stochastique ou de jeux à somme nulle.

Lorsque l'espace d'état et le temps sont continus, la programmation dynamique conduit à l'étude d'une équation aux dérivées partielles de type Hamilton-Jacobi-Bellman. Certaines des études décrites précédemment ont pu être abordées directement sur l'équation, au moyen de la notion de solution de viscosité. Nous avons aussi étudié la résolution de cette équation au moyen d'algorithmes à base d'itérations sur les politiques et de méthodes multigrilles. L'ensemble de ces travaux ont permis la résolution de problèmes de gestion de portefeuille.

Les approches décrites ci-dessus apparaissent aussi dans d'autres domaines que le contrôle optimal. Par exemple, l'algèbre max-plus peut se voir comme la limite de déformations de l'algèbre usuelle. Nous avons ainsi obtenu des résultats de perturbations de valeurs propres de matrices, ainsi que des caractérisations de principes de grandes déviations à la loi des grands nombres. Nous avons par ailleurs étudié certains systèmes à retard ou le rang des pages du Web, en faisant appel en partie à des techniques d'applications monotones contractantes au sens large.

Abstract

The main part of this memoir is related to the study of deterministic or stochastic optimal control problems by the dynamic programming method.

Max-plus or tropical algebra allows one to think of dynamic programming operators associated to undiscounted deterministic optimal control problems as linear maps. This led us to studying some linear systems and spectral problems over the max-plus algebra, to develop a theory of idempotent probabilities and a max-plus finite element method.

More generally, dynamic programming operators are monotone maps that are non-expansive in the L^∞ norm. We have obtained several results in Perron-Frobenius theory (concerning the spectral problem, or asymptotics of iterates of these maps), and applied them to the study of stochastic optimal control problems and zero-sum games.

When the state space and time are continuous, dynamic programming leads to a Hamilton-Jacobi-Bellman type partial differential equation. Some of the studies described above have been done directly on this equation, with the help of viscosity solutions techniques. This equation has also been studied numerically by using algorithms based on policy iterations and multigrid methods. This work has been used to solve some portfolio selection problems.

The above approaches also appear in other domains than optimal control. For instance, max-plus algebra can be seen as the limit of deformations of usual algebra. This point of view allowed us to obtain matrix eigenvalues perturbations results, together with characterisations of large deviation principles. We also applied techniques of monotone, non-expansive maps to the study of some delay systems or of the rank of Web pages.

Table des matières

Introduction

La plupart des travaux présentés ici concernent ou sont reliés à l'étude de problèmes de contrôle optimal déterministe ou stochastique, voire de jeux à somme nulle, par la méthode de la programmation dynamique. Ces problèmes sont abordés sous les trois approches de l'algèbre max-plus, des fonctions monotones contractantes au sens large, et des équations aux dérivées partielles. Nous précisons ci-dessous les motivations de ces approches issues du contrôle optimal.

L'algèbre max-plus

Considérons le problème le plus simple de contrôle optimal déterministe non-actualisé à temps et espaces discrets et à horizon fini $k \in \mathbb{N}^*$:

$$v_i(k) = \max \sum_{s=0}^{k-1} \ell(\mathbf{x}_s, \mathbf{x}_{s+1}) + \phi(\mathbf{x}_k) \ , \tag{1}$$

où le maximum est pris sur toutes les suites $(\mathbf{x}_s)_{s \geq 0}$ à valeurs dans l'ensemble fini $\Omega = \{1, \dots, n\}$, et partant de $\mathbf{x}_0 = i$. Ici $\ell(i,j)$ désigne le gain pour aller de l'état i à l'état j, où l'on pose $\ell(i,j) = -\infty$ lorsqu'on ne peut pas aller de i à j, et $\phi(i)$ désigne le gain final dans la position i. L'équation classique de la programmation dynamique s'écrit :

$$v_i(k) = \max_{1 \leq j \leq n} (\ell(i,j) + v_j(k-1)) \ , \tag{2}$$

avec la condition initiale $v_i(0) = \phi(i)$, $i \in \Omega$. Si dans $\mathbb{R} \cup \{-\infty\}$, on note $+\!\!\!\!+$ l'opération $(a,b) \mapsto \max(a,b)$ et $\times\!\!\!\!\times$ l'addition usuelle, et si on construit la matrice $B := (\ell(i,j))_{i,j=1,\dots,n}$, alors on reconnaît dans (2) le produit matriciel

$$v(k) = Bv(k-1) \ , \tag{3}$$

où pour tout $k \geq 0$, $v(k)$ désigne le vecteur de $(\mathbb{R} \cup \{-\infty\})^n$ de coordonnées $(v_i(k))_{1 \leq i \leq n}$, qui représente donc la fonction valeur du problème. L'équation de la programmation dynamique est donc linéaire dans l'*algèbre max-plus* qui est précisément le semi-corps idempotent obtenu en considérant l'ensemble $\mathbb{R} \cup \{-\infty\}$ muni de l'addition $(a,b) \mapsto a +\!\!\!\!+ b = \max(a,b)$ et de la multiplication $(a,b) \mapsto a \times\!\!\!\!\times b = a + b$.

Si l'on avait considéré un problème de minimisation de coût, au lieu de la maximisation de gain, on aurait eu besoin, au lieu de l'algèbre max-plus, de l'*algèbre min-plus* qui est le semi-corps idempotent, isomorphe au précédent, obtenu en considérant l'ensemble $\mathbb{R} \cup \{+\infty\}$ muni des lois $a +\!\!\!\!+ b = \min(a,b)$ et $a \times\!\!\!\!\times b = a + b$.

Considérons maintenant le problème de contrôle déterministe non actualisé, à temps et espace continus et à horizon fini $t > 0$, dont le problème du calcul des variations est un cas particulier (pour $\dot{\mathbf{x}} = \mathbf{u}$) :

$$v(t,x) = \sup \int_0^t \ell(\mathbf{x}(s), \mathbf{u}(s))\, ds + \phi(\mathbf{x}(t)) \ , \tag{4a}$$

où le supremum est pris sur toutes les trajectoires $(\mathbf{x}(\cdot), \mathbf{u}(\cdot))$ à valeurs dans $\Omega \times U$, où $\Omega \subset \mathbb{R}^n$ est l'espace des *états*, et $U \subset \mathbb{R}^m$ est l'ensemble des *contrôles*, et vérifiant

$$\dot{\mathbf{x}}(s) = g(\mathbf{x}(s), \mathbf{u}(s)), \text{ pour tout } s \in [0,t] \text{ et } \mathbf{x}(0) = x \ , \tag{4b}$$

et où la *condition initiale* $x \in \Omega$ est donnée.

Ce problème peut être vu comme une version continue de (1), ce qui permet de voir l'équation d'Hamilton-Jacobi que vérifie v,

$$\frac{\partial v}{\partial t} - H(x, \frac{\partial v}{\partial x}) = 0, \quad x \in \Omega, t > 0 \quad \text{ où } H(x,p) = \sup_{u \in U}(p \cdot g(x,u) + \ell(x,u)) \ , \tag{5a}$$

$$v(0,x) = \phi(x) \ , \tag{5b}$$

comme une équation max-plus linéaire. En particulier, comme l'a remarqué Maslov [Mas73], les solutions de (5a) vérifient un principe de superposition max-plus : si v et w sont deux solutions, et si $\lambda, \mu \in \mathbb{R}$, $\sup(\lambda + v, \mu + w)$ est encore solution de (5a).

Les structures de type max-plus, appelées parfois semi-anneaux idempotents, parfois dioïdes, parfois "max-algebras", parfois algèbres tropicales,..., ont été introduites et développées indépendamment par plusieurs écoles, en relation avec plusieurs domaines. En particulier dans le contexte du contrôle optimal et des équations aux dérivées partielles d'Hamilton-Jacobi, on peut citer les travaux de l'école russe [Rom67, MS92, KM97, Sam02, LMS01], mais aussi [Qua90, Gon96, GM02]. Mais ces structures sont d'abord apparues avec la théorie des treillis [DJLC53]. Puis les motivations sont venues de la Recherche Opérationnelle [CG79, GM79, Vor67, Zim76, Zim81] : problèmes de plus court chemin, problèmes d'ordonnancement, optimisation discrète. Dans l'étude des systèmes à événements discrets, et en particulier des systèmes de production, l'algèbre max-plus permet de représenter de manière linéaire les phénomènes de synchronisation [BCOQ92, HOvdW05]. En informatique théorique, le semi-anneau tropical, qui désigne alors l'ensemble $\mathbb{N} \cup \{+\infty\}$ muni des lois $a \oplus b = \min(a,b)$ et $a \otimes b = a + b$, a permis la résolution de problèmes de décision en théorie des automates. Comme nous le verrons plus loin, l'algèbre max-plus peut être vue comme la limite de déformations de l'algèbre usuelle. Elle apparaît donc naturellement dans les asymptotiques de type WKB [Mas73, MS92, KM97], les grandes déviations [Puh01], les asymptotiques à température nulle en physique statistique [CG86] ou les perturbations de valeurs propres (que nous présentons au chapitre 5). L'algèbre max-plus est apparue récemment en géométrie algébrique [GKZ94, Vir01, Mik04, SS04] sous les nom de géométrie tropicale, et en théorie des représentations sous le nom de combinatoire tropicale.

La linéarité de l'équation de la programmation dynamique de problèmes de contrôle optimal déterministe a inspiré plusieurs de nos travaux.

L'équation (3) peut se voir comme l'analogue de l'équation de Kolmogorov associée à une chaîne de Markov. Dans le chapitre 3, nous présentons une théorie des probabilités idempotentes et des processus de Bellman (analogues aux processus de Markov) que l'on peut voir comme une vision probabiliste de l'optimisation et du contrôle optimal.

Le principe de superposition pour les solutions d'équations d'Hamilton-Jacobi a motivé le développement d'une méthode des éléments finis max-plus, que nous présentons dans la section 4.2.

Certains problèmes d'algèbre linéaire min-plus ou max-plus permettent de résoudre des problèmes de contrôle optimal. Par exemple, l'équation $v = Bv + c$ apparaît dans la résolution de problèmes de contrôle à temps d'arrêt optimal. Elle se résout classiquement à l'aide de l'étoile de Kleene.

Le problème de contrôle optimal déterministe ergodique, i.e. le problème de la maximisation du gain moyen en temps long :

$$\lambda = \sup \left(\limsup_{k\to+\infty} \frac{1}{k} \sum_{s=0}^{k-1} \ell(\mathbf{x}_s, \mathbf{x}_{s+1}) \right) , \tag{6}$$

est associé à la résolution du problème spectral $Bv = \lambda v$. En effet la solution λ de (6) est valeur propre de la matrice B définie plus haut, et les politiques optimales stationnaires de ce problème sont fournies par les contrôles optimaux pour les vecteurs propres v associés à λ, qui peuvent eux-mêmes être considérés comme des “prix” induisant ces politiques stationnaires. Le problème spectral max-plus a été beaucoup étudié dans le cadre de la dimension finie dans les premiers travaux en algèbre max-plus [CG79, Rom67, Vor67, GM77, CDQV83]. Il a été moins étudié dans le cas de la dimension infinie et peu étudié dans le cas d'un espace d'état non compact. C'est ce que nous faisons dans le chapitre 2.

Le problème $Bv = c$ apparaît dans l'identification du coût terminal à partir de la fonction valeur. Il permet aussi de caractériser la limite de problèmes de contrôle optimal qui dépendent d'un paramètre (voir la section 3.4). Dans la section 1.1, nous présentons des résultats d'existence et d'unicité de la solution v en termes de sous-différentiels généralisés, ce qui étend au cas de la dimension infinie des résultats de Vorobyev [Vor67] et Zimmermann [Zim76].

D'autres problèmes d'algèbre linéaire max-plus peuvent avoir des motivations autres que le contrôle optimal. Dans la section 1.3, nous présentons des résultats sur la symétrisation de l'algèbre max-plus, permettant de résoudre des systèmes linéaires du type $Ax + b = Cx + d$. Ces travaux anciens se trouvent être reliés à des résultats récents en géométrie tropicale. Dans la section 1.2, nous présentons des résultats de séparation par des formes linéaires sur un semi-corps idempotent.

Comme nous l'avons dit plus haut, l'algèbre min-plus apparaît naturellement dans les asymptotiques de type WKB ou grandes déviations, parce qu'elle peut être vue comme la limite de déformations de l'algèbre usuelle. En effet :

$$\epsilon^a + \epsilon^b \asymp \epsilon^{\min(a,b)} , \qquad \epsilon^a \times \epsilon^b = \epsilon^{a+b} \quad \text{lorsque } \epsilon \to 0^+ . \tag{7}$$

Ainsi, le semi-anneau $\mathbb{R}_\epsilon$, qui est l'ensemble $\mathbb{R} \cup \{+\infty\}$, muni de l'addition $(a, b) \mapsto \frac{1}{\log \epsilon} \log(\epsilon^a + \epsilon^b)$ et de la multiplication $(a, b) \mapsto a + b$, est isomorphe au semi-corps $(\mathbb{R}_+, +, \times)$, mais lorsque ϵ tend vers 0^+, $\mathbb{R}_\epsilon$ tend vers le semi-anneau min-plus. Cette idée, introduite par Maslov [Mas73], permet d'utiliser des résultats algébriques pour résoudre des problèmes d'asymptotiques. C'est ce que nous faisons dans la section 3.4, où nous présentons l'application des résultats de la section 1.1 sur l'unicité de la solution des équations linéaires de type $Bv = c$ à l'identification de taux de grandes déviations. De même, dans le chapitre 5, nous appliquons ces idées aux perturbations singulières de valeurs propres.

Applications monotones homogènes

Considérons maintenant le problème de contrôle optimal stochastique :

$$v_i(k) = \sup \mathbb{E}\left[\sum_{s=0}^{k-1} \ell(\mathbf{x}_s, \mathbf{u}_s) + \phi(\mathbf{x}_k)\right] , \tag{8}$$

où le supremum est pris sur toutes les chaînes de Markov contrôlées $(\mathbf{x}_s)_{s\geq 0}$ à valeurs dans l'ensemble fini $\Omega = \{1,\dots,n\}$, partant de $\mathbf{x}_0 = i$ et de probabilité de transition de l'état $\mathbf{x}_s = i$ à l'état $\mathbf{x}_{s+1} = j$ égale à $P_{ij}^{\mathbf{u}_s}$, et toutes les stratégies $(\mathbf{u}_s)_{s\geq 0}$ à valeurs dans l'espace des contrôles U. On note f l'application sur $\mathbb{R}^n$, telle que

$$f_i(v) = \sup_{u\in U}\left(\ell(i,u) + \sum_{j=1}^{n} P_{ij}^u v_j\right) ,$$

qui est à valeurs finies dès que ℓ est bornée supérieurement, et $U \neq \emptyset$. On obtient que l'équation de la programmation dynamique associée à (8) s'écrit :

$$v(k) = f(v(k-1)) , \tag{9}$$

avec la condition initiale $v(0) = w$, où pour tout $k \geq 0$, $v(k)$ est le vecteur de $\mathbb{R}^n$ de coordonnées $(v_i(k))_{1\leq i\leq n}$, qui représente la fonction valeur du problème, et où w est le vecteur de coordonnées $(\phi(i))_{1\leq i\leq n}$. L'application f est max-plus linéaire si les probabilités de transitions sont dégénérées (cas déterministe). Elle est affine si l'ensemble U est réduit à un élément (chaîne de Markov non contrôlée). En général elle a les propriétés de monotonie (M) et d'homogénéité additive (AH) suivantes

$$\begin{aligned} &\text{(M)} && x \leq y \;\Rightarrow\; f(x) \leq f(y) , \\ &\text{(AH)} && f(\lambda + x) = \lambda + f(x) , \end{aligned}$$

pour tous $x, y \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$, où $\leq$ est l'ordre partiel de $\mathbb{R}^n$ et $\lambda + (x_1,\dots,x_n) := (\lambda + x_1,\dots,\lambda + x_n)$. En outre, elle est contractante au sens large pour la norme infinie :

$$\text{(C)} \quad \|f(x) - f(y)\|_\infty \;\leq\; \|x - y\|_\infty \quad \forall x, y \in \mathbb{R}^n .$$

Elle est de plus *convexe* dans le sens que toutes ses fonctions coordonnées $f_i : \mathbb{R}^n \to \mathbb{R}$ sont convexes. Les fonctions monotones homogènes sont parfois appelées *fonctions topicales*. Si on rajoute un facteur d'actualisation au problème, l'opérateur de la programmation dynamique n'est plus additivement homogène, mais il conserve la propriété de sous-homogénéité additive suivante :

$$\text{(ASH)} \quad f(\lambda + x) \leq \lambda + f(x) ,$$

pour $\lambda \geq 0$ et $x \in \mathbb{R}^n$. Notons qu'inversement toute application monotone sous-homogène convexe de $\mathbb{R}^n$ dans lui-même correspond à l'opérateur de la programmation dynamique d'un problème de contrôle optimal stochastique (ce résultat est une conséquence élémentaire de la dualité de Legendre-Fenchel).

Si maintenant on considérait un problème de jeux à deux joueurs et à somme nulle, on obtiendrait une équation de la programmation dynamique de la forme (9) pour une application monotone et homogène ou sous-homogène mais pas forcément convexe, et inversement toute application monotone sous-homogène de $\mathbb{R}^n$ dans lui-même correspond à l'opérateur de la programmation dynamique d'un jeux à deux joueurs et somme nulle, qui peut être choisi déterministe (mais l'espace des actions est infini). Ce point de vue sur les jeux a été utilisé dans [RS01, NS03].

On peut bien sûr généraliser les propriétés précédentes au cas d'un espace d'état Ω infini : on se retrouve alors avec une application monotone et contractante au sens large sur l'espace des fonctions bornées. Dans le cas de diffusions contrôlées, l'équation de la programmation dynamique est une équation d'Hamilton-Jacobi-Bellman :

$$\frac{\partial v}{\partial t} - H(x, \frac{\partial v}{\partial x}, \frac{\partial^2 v}{\partial x^2}) = 0, \quad x \in \Omega \tag{10a}$$

$$\text{où } H(x,p,X) = \sup_{u \in U}(\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(x,u)X_{ij} + p \cdot g(x,u) + \ell(x,u)) \;, \tag{10b}$$

où, pour tout (x, u), la matrice $(a_{ij}(x, u))$ est symétrique positive semi-définie. Le semi-groupe d'évolution est alors formé d'applications non-linéaires monotones sous-homogènes.

Les problèmes classiques de contrôle optimal stochastique ou de jeux conduisent chacun à un problème particulier pour l'opérateur de la programmation dynamique f qui lui est associé. Par exemple un problème à horizon infini actualisé ou non, ou à temps d'arrêt optimal conduit à l'équation de la programmation dynamique $v = f(v)$, qui est donc une équation de point fixe. L'algorithme classique d'itération sur les valeurs appliqué à cette équation consiste alors à utiliser l'algorithme de point fixe (9). Sa convergence dépend des propriétés de contraction éventuelles de f. Si l'on considère un problème de contrôle ergodique, la valeur λ du problème s'interprète comme une valeur propre (additive) de f et les stratégies optimales stationnaires sont déterminées par les vecteurs propres v pour cette valeur propre, i.e. les solutions v de $f(v) = \lambda + v$. En particulier, la structure de l'ensemble des vecteurs propres permet de déterminer la structure de l'ensemble des stratégies optimales. Plus généralement, la valeur d'un problème de contrôle ergodique est liée à la limite des suites $f^k(v)/k$ lorsque k tend vers l'infini.

L'application exponentielle terme à terme $\exp : \mathbb{R}^n \to (\mathbb{R}_+^*)^n,\ x \mapsto (\exp(x_1), \ldots, \exp(x_n))$ envoie par conjugaison les applications monotones additivement homogènes vers des applications monotones (positivement) homogènes (i.e. vérifiant $f(\lambda x) = \lambda f(x)$ pour tout $\lambda > 0$) sur le cône positif de $\mathbb{R}^n$. Les problèmes relatifs aux applications monotones additivement homogènes soulevés plus haut par les problèmes de contrôle optimal peuvent donc être englobés dans la théorie de Perron-Frobenius non-linéaire, qui traite des applications non linéaires sur des cônes, vérifiant des conditions de monotonie, de contraction ou d'homogénéité [KR50, Nus88, MPN02]. Le chapitre 6 traite de quelques uns de ces problèmes.

Le chapitre 7 est consacré à des travaux indépendants faisant appel en partie à des techniques de fonctions monotones contractantes au sens large.

Équations aux dérivées partielles

Comme nous l'avons vu plus haut, l'équation de la programmation dynamique associée à un problème de contrôle optimal déterministe ou stochastique à temps et espace continus est une équation d'Hamilton-Jacobi ou d'Hamilton-Jacobi-Bellman, les équations d'évolution (5a) et (10) correspondant au cas de problèmes à horizon fini. Un problème de jeux à somme nulle conduit quant à lui à une équation d'Isaacs de même forme que (5a) et (10), selon que le problème est déterministe ou stochastique, mais avec un hamiltonien H qui n'est ni convexe, ni concave. De plus, selon le modèle (utilisation de contrôles singuliers ou impulsionnels), ces équations prennent la forme d'une inéquation variationnelle, ou quasi-variationnelle.

Afin d'obtenir de l'information sur un problème de contrôle optimal particulier, on peut alors être amené à dégager des propriétés de l'équation aux dérivées partielles associée : structure de l'ensemble des solutions, convergence de perturbations ou d'approximations numériques de ces équations, etc. De plus, une étape importante est la résolution numérique de ces équations. Les travaux présentés dans la partie III relèvent de ces problématiques et font appel uniquement à des techniques d'équations aux dérivées partielles : discrétisations de type différences finies, estimations d'erreurs numériques pour des normes de type Sobolev, existence et unicité de solutions par des techniques de solutions de viscosité.

Le chapitre 8 rappelle rapidement mes travaux de thèse. Ceux-ci concernaient le développement d'un algorithme de résolution rapide d'équations d'Hamilton-Jacobi-Bellman discrétisées par différences finies. L'algorithme développé est basé sur l'algorithme de Howard ou d'itérations sur les politiques et les méthodes multigrilles.

Le chapitre 9 est consacré à l'étude théorique et numérique de quelques problèmes de gestion de portefeuille.

Organisation du mémoire

Nous avons classé les travaux par approche et par thème. L'ordre des chapitres et sections n'est donc pas chronologique. Chacune des trois parties contient les travaux de chacune des trois approches décrites ci-dessus, qu'ils soient ou non liés à des problèmes de contrôle optimal.

Les références bibliographiques se trouvent à la fin de ce mémoire. Les publications personnelles, de la forme [J1], [B5],... sont listées en premier. Les autres publications, de la forme [CG79, BCOQ92],..., sont listées ensuite.

Les publications et pré-publications [J3, J6, J7, J8, J10, J12, J14, J15, J16, B5, B6, B7, C3, C18, S1, S2] sont assez représentatives des travaux présentés dans ce mémoire. Elles ont été regroupées avec [T1, J5, J9, J13, B2, B8, B9, C7, C17, C19, R2] dans l'archive compressée (26M) disponible provisoirement dans :
`http ://minimal.inria.fr/publis-akian-non-distribuees/publis-akian-hdr.tar.gz`.

L'ensemble de ces publications est suffisant pour obtenir preuves et détails des travaux présentés dans ce mémoire.

Préliminaires et notations

On dit qu'un monoïde $(\mathbb{D}, \oplus)$ (resp. un semi-anneau ou un semi-corps $(\mathbb{D}, \oplus, \otimes)$) est *idempotent* si $a \oplus a = a$ pour tout $a \in \mathbb{D}$. Rappelons que l'algèbre max-plus, que l'on notera $\mathbb{R}_{\max}$, est le semi-corps idempotent formé de l'ensemble $\mathbb{R} \cup \{-\infty\}$ muni de l'addition $(a, b) \mapsto a \oplus b = \max(a, b)$ et de la multiplication $(a, b) \mapsto a \otimes b = a + b$. Son élément neutre pour l'addition, ou *zéro*, est $-\infty$, son élément neutre pour la multiplication, ou *unité*, est 0. Dans la suite, nous utiliserons respectivement les notations $\oplus$, $\bigoplus$, $\otimes$ (ou la concaténation) $\bigotimes$, $\mathbb{0}$ et $\mathbb{1}$ pour l'addition, la somme, la multiplication, le produit, le zéro et l'unité de $\mathbb{R}_{\max}$, ou de tout autre semi-anneau idempotent. De manière générale, on utilisera des fontes avec doubles barres pour désigner des opérations max-plus. Ainsi, si a et b sont dans $\mathbb{R}_{\max}$, $a \oplus b$ désigne $\max(a, b)$, $a \otimes b$ ou ab désigne la somme usuelle $a + b$, $a \oslash b$ désigne la différence usuelle $a - b$ lorsque $b \neq \mathbb{0}$, et a^n désigne $n \times a$.

L'algèbre min-plus sera notée $\mathbb{R}_{\min}$. On aura aussi besoin parfois du *semi-anneau max-plus complet*, noté $\overline{\mathbb{R}}_{\max}$, qui est formé de l'ensemble $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ muni de l'addition $(a, b) \mapsto \max(a, b)$ et de la multiplication $(a, b) \mapsto a + b$, avec la convention : $-\infty + (+\infty) = +\infty + (-\infty) = -\infty$. On définit de manière analogue le *semi-anneau min-plus complet*, $\overline{\mathbb{R}}_{\min}$.

La plupart des notions algébriques classiques ont des analogues max-plus (ou min-plus, ou encore sur un semi-anneau $\mathbb{D}$), et sauf contre-indication, nous utiliserons les même notations. Par exemple, $\mathbb{D}^n$ est l'ensemble des vecteurs de dimension n et $\mathbb{D}^{n \times p}$ est l'ensemble des matrices de taille $n \times p$ à coordonnées dans $\mathbb{D}$. Ces ensembles sont munis des opérations vectorielles et matricielles définies de manière usuelle à partir des opérations sur les scalaires, et notées de la même façon. La *matrice nulle* de taille $n \times p$, $\mathbb{0}_{np}$, aussi notée $\mathbb{0}$, a toutes ses coordonnées égales au zéro de $\mathbb{D}$. La *matrice identité* de taille $n \times n$, $\mathbb{I}_n$, aussi notée $\mathbb{I}$, a ses éléments diagonaux égaux à l'unité de $\mathbb{D}$, et ses éléments hors diagonaux égaux au zéro de $\mathbb{D}$. Étant donnée une matrice $A = (A_{ij}) \in \mathbb{D}^{n \times p}$, on notera $A_{i\cdot}$ et $A_{\cdot j}$ la i-ième ligne et la j-ième colonne de A. On notera aussi A l'application linéaire $\mathbb{D}^p \to \mathbb{D}^n$ qui à un vecteur x associe le vecteur Ax. On définit les *semi-modules* et les *sous-semi-modules* sur $\mathbb{R}_{\max}$ ou $\mathbb{D}$ de la même manière que les modules et sous-modules sur des anneaux. Étant donné un sous-ensemble F d'un $\mathbb{D}$-semi-module M, on dit que F *génère* M, ou est une *famille génératrice* de M, si tout élément $x \in M$ peut s'écrire comme combinaison linéaire finie des éléments de F, i.e. : $x = \bigoplus_{f \in F} \lambda_f . f$, où $(\lambda_f)_{f \in F}$ est une famille d'éléments de $\mathbb{R}_{\max}$ telle que $\lambda_f = \mathbb{0}$ sauf pour un ensemble fini d'éléments f de F. Un semi-module est dit *finiment engendré* s'il possède une famille génératrice finie.

Étant donné un monoïde idempotent $(\mathbb{D}, \oplus)$, on lui associe un *ordre naturel* $a \leq b \Leftrightarrow a \oplus b = b$. L'ensemble $\mathbb{D}$ muni de cet ordre naturel est alors un *sup-semi-treillis*, c'est-à-dire que tout sous-ensemble fini de $\mathbb{D}$ admet un plus petit majorant ou supremum. De plus, l'élément neutre de $\mathbb{D}$ est aussi son plus petit élément pour l'ordre naturel. Inversement tout sup-semi-treillis $\mathbb{D}$ ayant un plus petit élément est un monoïde idempotent pour la loi $a \oplus b = a \vee b$, d'élément neutre son plus petit élément. Ici et plus loin, étant donné un ensemble partiellement ordonné $(\mathbb{D}, \leq)$, nous noterons sup ou $\vee$ (resp. inf ou $\wedge$) le supremum (resp. l'infimum ou plus grand minorant) lorsqu'il existe. Tout semi-anneau idempotent et tout semi-module sur un tel semi-anneau (qui est alors nécessairement un monoïde idempotent) est donc muni d'un ordre naturel. Par exemple sur $\mathbb{R}_{\max}$ et $\overline{\mathbb{R}}_{\max}$ l'ordre naturel est identique à l'ordre usuel de $\overline{\mathbb{R}}$, sur les $\mathbb{R}_{\max}$-semi-modules de vecteurs ou de matrices sur $\mathbb{R}_{\max}$ l'ordre naturel est identique à l'ordre produit, sur $\mathbb{R}_{\min}$ et $\overline{\mathbb{R}}_{\min}$ l'ordre naturel est identique à l'ordre opposé à l'ordre usuel de $\overline{\mathbb{R}}$. Finalement, toute application linéaire d'un semi-module idempotent dans un autre, préserve nécessairement l'ordre naturel.

L'existence de la relation d'ordre naturel nous amène à utiliser des notions introduites en théorie des treillis, et en particulier la notion de résiduation qui occupe une place importante en algèbre max-plus. Rappelons qu'un ensemble partiellement ordonné $(S, \leq)$ est un *treillis* si tout sous-ensemble fini non vide de S admet un infimum et un supremum. On dit que c'est un *treillis complet* si tout sous-ensemble (éventuellement vide) de S admet un infimum, ou de manière équivalente si tout sous-ensemble de S admet un supremum. Considérons deux ensembles partiellement ordonnés quelconques $(\mathscr{F}, \leq_{\mathscr{F}})$ et $(\mathscr{G}, \leq_{\mathscr{G}})$, et une application $B : \mathscr{F} \to \mathscr{G}$. On dit que B est *isotone* si elle préserve l'ordre, i.e. si $f \leq_{\mathscr{F}} f' \implies Bf \leq_{\mathscr{G}} Bf'$. On dit que B est *résiduable* [BJ72, BCOQ92] si il existe une application $C : \mathscr{G} \to \mathscr{F}$, telle que (B, C) satisfait aux conditions équivalentes suivantes :

$$I_{\mathscr{F}} \leq_{\mathscr{F}} CB, \quad I_{\mathscr{G}} \geq_{\mathscr{G}} BC, \quad \text{et } B, C \text{ sont des applications isotones,} \tag{11a}$$

$$(Bf \leq_{\mathscr{G}} g \iff f \leq_{\mathscr{F}} Cg) \qquad \forall f \in \mathscr{F},\ g \in \mathscr{G}\ , \tag{11b}$$

$$Cg = \max\nolimits_{\mathscr{F}}\{f \mid Bf \leq_{\mathscr{G}} g\} \qquad \forall g \in \mathscr{G}\ , \tag{11c}$$

$$Bf = \min\nolimits_{\mathscr{G}}\{g \mid f \leq_{\mathscr{F}} Cg\} \qquad \forall f \in \mathscr{F}\ , \tag{11d}$$

dans lesquelles $\max_{\mathscr{F}}$ (resp. $\min_{\mathscr{G}}$) désigne le maximum (resp. minimum) d'un ensemble pour l'ordre $\leq_{\mathscr{F}}$ (resp. $\leq_{\mathscr{G}}$), et $I_{\mathscr{A}}$ désigne l'application identité de l'ensemble $\mathscr{A}$. L'application C est appelée la résiduée de B. Elle est unique et est notée $B^\sharp$. De même $C : \mathscr{G} \to \mathscr{F}$ est dite *dualement résiduable* si il existe $B : \mathscr{F} \to \mathscr{G}$, telle que (B, C) satisfait aux conditions équivalentes (11).

Si B est résiduable, alors B est "infiniment" additive, i.e.

$$B\left(\sup\nolimits_{\mathscr{F}} F\right) = \sup\nolimits_{\mathscr{G}}\{Bf \mid f \in F\}\ , \tag{12}$$

pour tout sous-ensemble F de $\mathscr{F}$ qui admet un supremum. De plus si $\mathscr{F}$ est complet, la propriété (12) caractérise les applications B résiduables. Ceci explique l'utilité des applications résiduables en algèbre linéaire max-plus.

On dit qu'un semi-anneau idempotent $(\mathbb{D}, \oplus, \otimes)$ est *complet* s'il est complet pour l'ordre naturel et si, pour tout a, les applications multiplication $\mathbb{D} \to \mathbb{D}$, $x \mapsto a \otimes x$ et $x \mapsto x \otimes a$ sont résiduables. De même un $\mathbb{D}$-*semi-module complet* est un $\mathbb{D}$-semi-module qui est complet pour l'ordre naturel et tel que les applications multiplication sont résiduables. Le semi-anneau $\overline{\mathbb{R}}_{\max}$ et le $\overline{\mathbb{R}}_{\max}$-semi-module $\overline{\mathbb{R}}_{\max}^n$ sont complets. Lorsque $\mathscr{F} = \overline{\mathbb{R}}_{\max}^n$, $\mathscr{G} = \overline{\mathbb{R}}_{\max}^p$ et B est l'application linéaire associée à une matrice à coefficients dans $\overline{\mathbb{R}}_{\max}$, notée aussi B, alors B est résiduable, et $B^\sharp$ est une application min-plus linéaire de matrice $-B^T$, où B^T est la matrice transposée de B.

Ces propriétés permettent de définir la somme des éléments d'un sous-sensemble quelconque X de scalaires, vecteurs ou matrices à coordonnées dans $\overline{\mathbb{R}}_{\max}$ (ou un semi-anneau idempotent complet) comme le supremum de X. Ainsi, si $A \in \overline{\mathbb{R}}_{\max}^{n\times n}$, l'*étoile de Kleene* de A est la matrice $A^\star = \mathbb{I} \oplus A \oplus A^2 \oplus \cdots$.

Première partie

Algèbre max-plus, contrôle déterministe, et asymptotiques

Chapitre 1

Systèmes linéaires max-plus

Nous présentons ici trois travaux différents que nous avons regroupé parce qu'ils sont tous trois relatifs à l'étude de systèmes linéaires par rapport à l'algèbre max-plus ou min-plus. Les techniques utilisées dans chacun des trois travaux sont différentes. La section 1.1 traite de l'équation $Bf = g$ et fait appel à la notion de sous-différentiels généralisés introduite en convexité abstraite. La section 1.2 traite de la séparation d'un certain type de "convexes" par des formes min-plus linéaires, et fait appel à la notion de treillis continu et à ses topologies. Elle est aussi reliée aux travaux en convexité abstraite. La section 1.3 traite de l'équation $Ax \oplus b = Cx \oplus d$, et fait appel à des outils combinatoires.

1.1 L'équation $Bf = g$

Les travaux présentés ici ont été obtenus en collaboration avec Stéphane Gaubert (INRIA Rocquencourt), et Vassili Kolokoltsov (Warwick University). Ils ont fait l'objet des articles publiés [J11, B6].

Nous nous posons ici la question de l'existence et de l'unicité de la solution f à l'équation $Bf = g$, lorsque B est un opérateur linéaire max-plus résiduable. Nous considérerons plus généralement le cas où B est un opérateur résiduable quelconque, ou, ce qui est équivalent à un changement de signe près, le cas où B conduit à une correspondance de Galois. Nous nous intéressons particulièrement à la dimension infinie, mais dans la situation où B a un noyau, que nous caractérisons dans la section 1.1.1. Une première motivation de cette étude vient de certains problèmes de convergence en épigraphe en analyse convexe, de grandes déviations à la loi des grands nombre en probabilités, ou de certains problèmes de contrôle optimal sensible au risque, que nous traiterons dans la section 3.4. Une deuxième motivation vient de l'étude de problèmes d'affectation optimale, à laquelle nous travaillons en ce moment avec Stéphane Gaubert et Vassili Kolokoltsov. Une dernière motivation vient du problème de la reconstruction du coût final d'un problème de contrôle optimal, à partir de sa fonction valeur à un certain instant, problème qui a été étudié sous des conditions de convexité par Goebel et Rockafellar.

1.1.1 Représentation de correspondances de Galois fonctionnelles

Fixons X et Y deux ensembles. Une *conjugaison de Moreau*, associée au noyau $b : X \times Y \to \overline{\mathbb{R}}$, est une application $B : \mathscr{F} \to \mathscr{G}$, où $\mathscr{F}$ et $\mathscr{G}$ sont des sous-ensembles respectifs de $\overline{\mathbb{R}}^Y$ et $\overline{\mathbb{R}}^X$, telle que :

$$Bf(x) = \sup\{b(x,y) - f(y) \mid y \in Y\}, \quad \forall x \in X \ . \tag{1.1}$$

Ici $b(x,y) - f(y)$ est une abréviation pour $b(x,y) + (-f(y))$, où on utilise la convention que $-\infty$ est absorbant pour l'addition. Un exemple important de conjugaison de Moreau est fourni par la transformée de Legendre-Fenchel entre deux espaces vectoriels topologiques localement convexes séparés (e.v.t.l.c.s.) en dualité X et Y, laquelle a pour noyau $b(x,y) = \langle x, y\rangle$. Comme les fonctions convexes semi-continues inférieurement (s.c.i.) sur X sont les images par la transformée de Legendre-Fenchel de fonctions quelconques, les images des conjugaisons de Moreau constituent une généralisation de la notion de fonction convexe s.c.i. Ainsi, les conjugaisons de Moreau sont utilisées en dualité non-convexe, voir par exemple [Sin97].

On voit facilement que si $B : \overline{\mathbb{R}}^Y \to \overline{\mathbb{R}}^X$ est une conjugaison de Moreau, alors l'application $f \mapsto B(-f)$ est un opérateur max-plus linéaire "infiniment" additif, donc résiduable. Étant donné l'existence du noyau b vérifiant (1.1), nous disons de cette application qu'elle est un opérateur max-plus linéaire à noyau. Le théorème 1.1 ci-dessous montre que le caractère infiniment additif est équivalent sous certaines conditions à l'existence d'un noyau, ceci dans le cadre plus général des applications résiduables ou des correspondances de Galois.

Rappelons que les *correspondances de Galois* sont définies en inversant la relation d'ordre de $\mathscr{G}$, mais pas celle de $\mathscr{F}$ dans (11), voir par exemple [GHK$^+$80]. De même, les *correspondances de Galois duales* sont définies en inversant la relation d'ordre de $\mathscr{F}$, mais pas celle de $\mathscr{G}$ dans (11). Par exemple, on dit que (B,C) est une *correspondance de Galois duale* entre les ensembles partiellement ordonnés $(\mathscr{F}, \leq_{\mathscr{F}})$ et $(\mathscr{G}, \leq_{\mathscr{G}})$, si $B : \mathscr{F} \to \mathscr{G}$, $C : \mathscr{G} \to \mathscr{F}$ et :

$$(g \geq_{\mathscr{G}} Bf \iff f \geq_{\mathscr{F}} Cg) \qquad \forall f \in \mathscr{F},\ g \in \mathscr{G} \ .$$

Étant donné B, l'unique application C telle que (B,C) est une correspondance de Galois duale est notée B°. L'avantage de la notion de correspondance de Galois, par rapport à celle de résiduation, est sa symétrie : si (B,C) est une correspondance de Galois (duale), alors (C,B) l'est aussi. Ainsi $(B^\circ)^\circ = B$.

Les propriétés énoncées dans le cadre de la résiduation (voir les préliminaires) montrent en particulier qu'une conjugaison de Moreau de $\overline{\mathbb{R}}^Y$ dans $\overline{\mathbb{R}}^X$ conduit à une correspondance de Galois duale.

Étant donné un *treillis* $(S, \leq)$ et un ensemble Y, nous disons qu'un sous-ensemble $\mathscr{F}$ de S^Y est un *treillis de fonctions* si c'est un sous-treillis du treillis S^Y muni de l'ordre produit (noté encore $\leq$). Une correspondance de Galois duale entre treillis de fonctions sera dite *fonctionnelle*, et on omettra le terme "duale". Si S admet un maximum $\top_S$, la *fonction de Dirac* au point $y \in Y$ et de valeur $s \in S$, notée δ_y^s, est l'application $\delta_y^s \in S^Y$ telle que $\delta_y^s(y') = s$ si $y' = y$, et $\delta_y^s(y') = \top_S$ sinon.

Théorème 1.1 ([B6, Theorem 2.1]). *Soient S, T deux treillis avec maximum, X, Y deux ensembles non vides, et soit $\mathscr{F} \subset S^Y$ (resp. $\mathscr{G} \subset T^X$) un treillis de fonctions contenant toutes les fonctions de Dirac de S^Y (resp. T^X). Alors (B, B°) est une correspondance de Galois duale entre $\mathscr{F}$ et $\mathscr{G}$ si et seulement si il existe deux applications $b : X \times Y \times S \to T$ et $b^\circ : Y \times X \times T \to S$ telles que : pour tout $(x,y) \in X \times Y$, $(b(x,y,\cdot), b^\circ(y,x,\cdot))$ est une correspondance de Galois duale entre S et T ; pour tout $(x,t) \in X \times T$, $b^\circ(\cdot, x, t) \in \mathscr{F}$; pour tout $(y,s) \in Y \times S$, $b(\cdot, y, s) \in \mathscr{G}$; et*

$$Bf = \sup\nolimits_{\mathscr{G}}\{b(\cdot, y, f(y)) \mid y \in Y\}, \quad \forall f \in \mathscr{F} \ , \tag{1.2a}$$

$$B^\circ g = \sup\nolimits_{\mathscr{F}}\{b^\circ(\cdot, x, g(x)) \mid x \in X\}, \quad \forall g \in \mathscr{G} \ . \tag{1.2b}$$

Dans ce cas, les applications b et b° sont déterminées de manière unique par (B, B°), puisque $b(\cdot, y, s) = B\delta_y^s$ et $b^\circ(\cdot, x, t) = B^\circ\delta_x^t$ pour tous $s \in S, t \in T, x \in X$ et $y \in Y$.

Le théorème 1.1 est inspiré du "théorème de représentation de Riesz" de Maslov et Kolokoltsov (voir par exemple [KM87] et [KM97, Theorem 1.4]) qui s'applique aux applications continues finiment additives B entre des treillis (non-complets) $\mathscr{F}$ et $\mathscr{G}$ de fonctions continues à valeurs réelles. Martínez-Legaz et Singer ont obtenu dans [MLS90, Theorems 3.1 et 3.5] (voir aussi [Sin97, Theorem 7.3 et la page 419]) les même conclusions que dans le théorème 1.1, dans le cas particulier où $\mathscr{F} = S^Y$, $\mathscr{G} = T^X$ et où S et T sont des treillis complets. On verra dans la section 3.1, un résultat similaire au théorème 1.1, mais sous une hypothèse plus faible d'additivité dénombrable.

Si Y est un espace topologique séparé et $S = T = \overline{\mathbb{R}}$, on peut appliquer le théorème 1.1 à l'ensemble $\mathscr{F} = \mathrm{sci}(Y, \overline{\mathbb{R}})$ des fonctions s.c.i. Si B est une conjugaison de Moreau de noyau $\bar{b}$, on a nécessairement $b(x, y, \lambda) = \bar{b}(x, y) - \lambda$ et $b^\circ(y, x, \lambda) = \bar{b}(x, y) - \lambda$.

1.1.2 Existence et unicité de solutions

On considère maintenant le problème de l'existence et de l'unicité de la solution f à l'équation $Bf = g$, lorsque (B, B°) est une correspondance de Galois fonctionnelle. Rappelons que par la théorie des correspondances de Galois, si $\mathscr{F}$ et $\mathscr{G}$ sont complets, alors

$$Bf = g \text{ a une solution } f \in \mathscr{F} \quad \Longleftrightarrow \quad BB^\circ g = g \; .$$

Ce résultat est très utile pour traiter de l'existence d'une solution, mais il ne permet pas d'aborder le problème de l'unicité, et nous cherchons plutôt ici une caractérisation qui généralise les résultats de Vorobyev [Vor67, Theorem 2.6], et Zimmermann [Zim76, Chapter 3], sur les opérateurs max-plus linéaires en dimension finie. Ceux-ci fournissent une condition combinatoire en termes de recouvrement de l'ensemble des états par une famille d'ensembles. Nous verrons que le cadre de la dimension infinie permet d'éclairer la signification géométrique (en termes de sous-différentiels) de la condition de Vorobyev et Zimmermann.

On se donne une correspondance de Galois fonctionnelle (B, B°) vérifiant les propriétés du théorème 1.1 avec $S = T = \overline{\mathbb{R}}$, X et Y des espaces topologiques séparés et $\mathscr{F} = \mathrm{sci}(Y, \overline{\mathbb{R}})$, $\mathscr{G} = \overline{\mathbb{R}}^X$. On note b et b° les noyaux de B et B° donnés par le théorème 1.1. Dans ce cas $\sup_{\mathscr{F}} = \sup$, et $b^\circ(\cdot, x, t)$ et $b(x, \cdot, s)$ sont s.c.i. (d'après le théorème 1.1 et [B6, Proposition 2.3]). De plus, les fonctions $b(x, y, \cdot)$ et $b^\circ(y, x, \cdot)$ de $\overline{\mathbb{R}}$ dans $\overline{\mathbb{R}}$ sont nécessairement décroissantes et continues à droite et envoient $-\infty$ sur $+\infty$. On suppose qu'il existe $\mathscr{S} \subset X \times Y$ tel que :

(H1) $\mathscr{S}_x = \{y \in Y \mid (x, y) \in \mathscr{S}\} \neq \emptyset$, pour tout $x \in X$;
(H2) $\mathscr{S}^y = \{x \in X \mid (x, y) \in \mathscr{S}\} \neq \emptyset$, pour tout $y \in Y$;
(H3) $b(x, y, \cdot)$ est une bijection $\overline{\mathbb{R}} \to \overline{\mathbb{R}}$ pour tout $(x, y) \in \mathscr{S}$;
(H4) $b(x, y, \cdot) \equiv -\infty$, lorsque $(x, y) \in X \times Y \setminus \mathscr{S}$.

Si B est une conjugaison de Moreau de noyau $\bar{b}$, les hypothèses (H1)–(H4) sont vérifiées si et seulement si $\bar{b}$ ne prend pas la valeur $+\infty$ et l'ensemble $\mathscr{S} := \{(x, y) \in X \times Y \mid \bar{b}(x, y) \in \mathbb{R}\}$ satisfait (H1)–(H2), ce qui est le cas pour la transformée de Legendre-Fenchel.

Sous les hypothèses précédentes, on définit la notion de sous-différentiel généralisé comme suit. Le *sous-différentiel* de $f \in \mathscr{F}$ au point $y \in Y$ par rapport à b (ou B), noté $\partial^b f(y)$, ou tout simplement $\partial f(y)$ est par définition l'ensemble :

$$\partial f(y) = \{x \in X \mid (x, y) \in \mathscr{S}, b(x, y', f(y')) \leq b(x, y, f(y)) \; \forall y' \in Y\} \; .$$

Le sous-différentiel de $g \in \mathscr{G}$ au point $x \in X$ par rapport à b°, $\partial^{b^\circ} g(x)$, est noté simplement $\partial^\circ g(x)$. Lorsque $b(x, y, \lambda) = \langle x, y\rangle - \lambda$, on retrouve la définition classique des sous-différentiels.

Plutôt que l'équation $Bf = g$, on traite le problème plus général suivant, étant donnés $g \in \mathscr{G}$ et $X' \subset X$:

$$(\mathscr{P}') : \qquad \text{trouver } f \in \mathscr{F} \text{ telle que } Bf \leq g \text{ et } Bf(x) = g(x) \text{ pour tout } x \in X' ,$$

qui est utile dans les applications aux grandes déviations de la section 3.4.

Existence de solutions

Les résultats d'existence et d'unicité sont basés sur les notions suivantes de recouvrement. Étant donnés deux ensembles Z et W, et Φ une application de Z dans l'ensemble $\mathcal{P}(W)$ de tous les sous-ensembles de W, on pose $\Phi^{-1}(w) = \{z \in Z \mid w \in \Phi(z)\}$ et $\mathrm{dom}(\Phi) := \{z \in Z \mid \Phi(z) \neq \emptyset\}$ (le *domaine* de Φ). Si $Z' \subset Z$ et $W' \subset W$, on dit que $\{\Phi(z)\}_{z\in Z'}$ est un *recouvrement* de W', si $\cup_{z\in Z'}\Phi(z) \supset W'$. Un élément $y \in Z'$ est dit *algébriquement essentiel* par rapport à ce recouvrement si $W' \not\subset \cup_{z\in Z'\setminus\{y\}}\Phi(z)$. Si Z est un espace topologique, y est dit *topologiquement essentiel* si, pour tout voisinage U de y dans Z', $W' \not\subset \cup_{z\in Z'\setminus U}\Phi(z)$. Finalement, le recouvrement de W' par $\{\Phi(z)\}_{z\in Z'}$ est dit *algébriquement* (resp. *topologiquement*) *minimal* si tous les éléments de Z' sont algébriquement (resp. topologiquement) essentiels.

Pour tout g d'un espace topologique Z dans $\overline{\mathbb{R}}$, on note $\mathrm{ldom}(g) := \{z \in Z \mid g(z) < +\infty\}$, $\mathrm{udom}(g) := \{z \in Z \mid g(z) > -\infty\}$, $\mathrm{dom}(g) := \mathrm{ldom}(g) \cap \mathrm{udom}(g)$ (le *domaine* de g), et $\mathrm{idom}(g) = \{z \in \mathrm{dom}(g) \mid \limsup_{z'\to z} g(z') < +\infty\}$.

La condition suffisante suivante pour l'existence d'une solution s'obtient de manière élémentaire.

Théorème 1.2 ([B6, Theorem 3.5, 1ère partie])**.** *Soient $X' \subset X$ et $g \in \mathscr{G}$. Considérons les assertions suivantes :*

(i) Le problème $(\mathscr{P}')$ a une solution ;

(ii) $X' \subset \mathrm{dom}(\partial^\circ g)$;

(iii) $\{(\partial^\circ g)^{-1}(y)\}_{y\in\mathrm{ldom}(B^\circ g)}$ est un recouvrement de $X' \cap \mathrm{udom}(g)$.

Alors (ii)⇔(iii)⇒(i).

Pour l'implication inverse dans le théorème 1.2, nous avons besoin de l'hypothèse technique suivante :

(H5) Les conditions (H6) ou (H6)$'$, et (H7) ou (H7)$'$ sont vérifiées, où

(H6) Y est discret ;

(H6)$'$ b est continue en la seconde variable, et $B^\circ g(y) > -\infty$ pour tout $y \in Y$;

(H7) $B^\circ g \in \mathscr{F}_c$;

(H7)$'$ b est coercive et $X' \subset \mathrm{idom}(g) \cup g^{-1}(-\infty)$.

Ici, on note $\mathscr{F}_c$ l'ensemble des $f \in \mathscr{F}$ telles que, pour tout $x \in X$, $y \mapsto b(x, y, f(y))$ a ses ensembles de sur-niveau finis relativement compacts, ce qui veut dire que pour tout $\beta \in \mathbb{R}$, l'ensemble $\{y \in Y \mid b(x, y, f(y)) \geq \beta\}$ est relativement compact. Lorsque Y est compact, $\mathscr{F}_c = \mathscr{F}$.

On dit que b est *coercif* si pour tous $\alpha \in \mathbb{R}$ et $x \in X$, et pour tout voisinage V de x dans X, la fonction

$$y \in Y \mapsto b^\alpha_{x,V}(y) = \sup_{z\in V} b(z, y, b^\circ(y, x, \alpha)) , \tag{1.3}$$

a ses ensembles de sous-niveau finis relativement compacts, ce qui veut dire que $\{y \in Y \mid b^{\alpha}_{x,V}(y) \leq \beta\}$ est relativement compact, pour tout $\beta \in \mathbb{R}$. Le noyau de la transformée de Legendre-Fenchel sur $\mathbb{R}^n$ est coercif.

L'hypothèse (H5) est toujours vérifiée lorsque Y est fini. Les conditions (H6) ou (H6)$'$ peuvent être vues comme une hypothèse de continuité, elles assurent que les ensembles $\{y \in Y \mid b(x, y, B^{\circ}g(y)) \geq \beta\}$ sont fermés. Les conditions (H7) ou (H7)$'$ peuvent être vues comme une hypothèse de compacité, elles assurent que ces même ensembles sont relativement compacts lorsque $x \in X'$.

Théorème 1.3 ([B6, Theorem 3.5, 2ième partie])**.** *Si* (H5) *est vérifiée, alors on a l'équivalence : (i)⇔(ii)⇔(iii) dans le théorème 1.2.*

Lorsque B est la transformée de Legendre-Fenchel sur $\mathbb{R}^n$, et g est une fonction convexe s.c.i. propre, le problème $(\mathscr{P}')$ a une solution pour tout sous-ensemble X' de X. L'implication (i)⇒(ii) dans le théorème 1.3 montre que g admet des sous-différentiels dans $\mathrm{idom}(g)$, ce qui est un résultat bien connu de convexité puisque $\mathrm{idom}(g)$ n'est autre que l'intérieur du domaine de g.

Unicité de solutions

On obtient une condition d'unicité en requiérant la minimalité du recouvrement qui intervient dans le résultat d'existence.

Théorème 1.4 ([B6, Theorem 4.6])**.** *Soient $X' \subset X$ et $g \in \mathscr{G}$. Supposons que $\{(\partial^{\circ} g)^{-1}(y)\}_{y \in \mathrm{ldom}(B^{\circ}g)}$ est un recouvrement de $X' \cap \mathrm{udom}(g)$, et notons Z_a (resp. Z_t) l'ensemble des éléments algébriquement (resp. topologiquement) essentiels par rapport à ce recouvrement. Posons $Z = Z_a \cup \mathrm{int}(Z_t)$, où l'on a noté $\mathrm{int}(Z_t)$ l'intérieur de Z_t relativement à $\mathrm{ldom}(B^{\circ}g)$. Supposons* (H5) *vérifiée, et supposons que $B^{\circ}g$ est quasi-continue dans son domaine. Alors, le problème $(\mathscr{P}')$ a une solution, et toute solution f de $(\mathscr{P}')$ vérifie*

$$f \geq B^{\circ}g, \quad \text{et} \quad f(y) = B^{\circ}g(y) \quad \text{pour tout } y \in Z \; .$$

Ici on dit qu'une fonction h d'un espace topologique Z à valeurs dans $\overline{\mathbb{R}}$ est *quasi-continue* si pour tout ouvert G de $\overline{\mathbb{R}}$, l'ensemble $h^{-1}(G)$ est inclus dans l'adhérence de son intérieur. Si h est s.c.i., cette condition est équivalente à celle que h est l'enveloppe s.c.i. de l'enveloppe semi-continue supérieure (s.c.s.) de h.

Théorème 1.5 ([B6, Theorem 4.7])**.** *Soient $X' \subset X$ et $g \in \mathscr{G}$. Considérons les assertions suivantes :*

(i) Le problème $(\mathscr{P}')$ a une unique solution,

(ii) $\{(\partial^{\circ} g)^{-1}(y)\}_{y \in \mathrm{ldom}(B^{\circ}g)}$ est un recouvrement topologiquement minimal de $X' \cap \mathrm{udom}(g)$.

Sous (H5)*, on a (i)⇒(ii). Si on suppose de plus que $B^{\circ}g$ est quasi-continue dans son domaine, alors (i)⇔(ii).*

La minimalité topologique du recouvrement dans la condition (ii) est une relaxation de la minimalité algébrique, qui est elle-même une condition de différentiabilité. Dans le cas de la transformée de Legendre-Fenchel, le fait qu'une fonction convexe s.c.i. propre g sur $\mathbb{R}^n$ soit essentiellement régulière est une propriété intermédiaire entre la minimalité topologique et la minimalité algébrique du recouvrement de (ii). Elle implique donc que si $f^{\star} \leq g$ et $f^{\star}(x) = g(x)$ pour tout $x \in \mathrm{idom}(g)$, alors $f = g^{\star}$ [B6, Corollary 6.4] (où $f^{\star}$ est la transformée de Legendre-Fenchel de f). En particulier

g a une unique pré-image par la transformée de Legendre-Fenchel. Cette propriété est sous-jacente à la preuve du théorème de Gärtner-Ellis sur les grandes déviations à la loi des grands nombres. Les résultats de cette section conduisent à des généralisations du théorème de Gärtner-Ellis dont nous présentons un aperçu dans la section 3.4.

1.2 Séparation et dualité quasi-convexe dans les groupes réticulés

Les travaux présentés ici ont été obtenus en collaboration avec Ivan Singer (Institut de Mathématiques, Bucarest). Ils ont fait l'objet de l'article publié [J13].

La séparation des ensembles convexes par des formes linéaires (théorème de Hahn-Banach) est un outil fondamental permettant de déduire de nombreuses propriétés des fonctions et ensembles convexes liées aux problèmes d'optimisation : images par la transformée de Fenchel-Legendre, conditions suffisantes d'extremum, dualité lagrangienne,... Une méthode pour étudier les problèmes d'optimisation non convexe consiste alors à généraliser la notion de convexité.

Celle-ci peut être généralisée de plusieurs manières. La transformée de Legendre-Fenchel étant un cas particulier de correspondance de Galois fonctionnelle, on peut définir les fonctions convexes généralisées comme les images par une correspondance de Galois fonctionnelle donnée, ce sont ces fonctions que nous avons étudié dans la section précédente.

De même, on peut considérer les ensembles convexes comme les images par une correspondance de Galois (duale) entre treillis d'ensembles, aussi appelée dualité en convexité abstraite [Sin97]. Si cette dualité est définie à partir d'une fonction de couplage $\varphi : X \times W \to \overline{\mathbb{R}}$, les ensembles convexes de X seront ceux que l'on peut “séparer inférieurement” par φ de tout point extérieur : C est un convexe de X si pour tout $y \in X \setminus C$, il existe $w \in W$ tel que $\varphi(x, w) \leq 0 < \varphi(y, w)$ pour tout $x \in C$. Si de plus cette fonction de couplage est la dualité entre un espace de Banach et son dual topologique à laquelle on ajoute une constante, on retrouve les ensembles convexes fermés classiques (par le théorème de séparation).

On peut aussi considérer les ensembles convexes comme étant les ensembles vérifiant les inégalités usuelles mais en se plaçant sur un semi-anneau quelconque au lieu de $\mathbb{R}$, par exemple en se plaçant sur l'algèbre max-plus. C'est ce qui est fait dans les travaux de Cohen, Gaubert, et Quadrat [CGQ04] et Cohen, Gaubert, Quadrat, et Singer [CGQS05], qui montrent des résultats de séparation d'ensembles max-plus convexes généraux par un couple de formes linéaires max-plus : si C est convexe max-plus et $y \notin C$ alors il existe deux formes affines max-plus φ, ψ telles que $\varphi(x) = \psi(x)$ pour tout $x \in C$ et $\varphi(y) > \psi(y)$.

Les résultats de Rubinov et Glover [RGJ95] sur la séparation des ensembles normaux fermés de $(\mathbb{R}_+^*)^n$ (où $\mathbb{R}_+^*$ est l'ensemble des réels strictement positifs) et ceux de Martínez-Legaz, Rubinov et Singer [MLRS02] sur la séparation des ensembles descendants (“downward”) fermés de $\mathbb{R}^n$ montrent que ces ensembles sont les ensembles convexes définis à partir d'une fonction de couplage qui n'est autre que le produit scalaire par rapport au semi-corps $(\mathbb{R}_+, \min, \times)$ ou *algèbre min-*$\times$ dans le premier cas et l'algèbre min-plus dans le second. Ces ensembles peuvent donc être vus comme des ensembles convexes min-$\times$ ou min-plus. Ils sont moins généraux que les ensembles convexes min-plus définis dans [CGQ04, CGQS05] car ils n'autorisent la séparation qu'avec une forme linéaire et dans une certaine direction, alors que ces derniers autorisent la séparation par un couple de formes linéaires. Il faut noter que la nécessité d'utiliser pour la séparation des convexes max-plus un couple de formes linéaires au lieu d'une seule forme linéaire vient de la non symétrisabilité de l'addition de l'algèbre max-plus ou min-plus (voir la section suivante).

Dans [J13], nous avons généralisé les résultats de séparation des ensembles normaux ou descendants fermés dans le cas où $\mathbb{R}_+^*$ et $\mathbb{R}$ sont remplacés par un groupe réticulé conditionnellement complet continu.

Rappelons qu'un *groupe réticulé conditionnellement complet* (g.r.c.c.) $\mathcal{R} = (\mathcal{R}, \leq, \divideontimes)$ est un ensemble $\mathcal{R}$ muni d'un ordre partiel $\leq$ tel que $(\mathcal{R}, \leq)$ est un *treillis conditionnellement complet* (i.e. tel que tout sous-sensemble non vide borné supérieurement de $\mathcal{R}$ admet un supremum), et d'une opération binaire $\divideontimes$ tel que $(\mathcal{R}, \divideontimes)$ est un groupe, dans lequel les multiplications par un élément sont isotones. Un tel groupe est nécessairement commutatif. De plus, en rajoutant un plus grand élément $\top$ et un plus petit élément $\perp$ à $\mathcal{R}$, on obtient un treillis complet $\overline{\mathcal{R}}$. En prolongeant la multiplication $\divideontimes$ par les lois $\underline{\divideontimes}$ et $\overline{\divideontimes}$ sur $\overline{\mathcal{R}}$, qui ne ne différent que par le fait que $\perp \underline{\divideontimes} \top = \perp$ et $\perp \overline{\divideontimes} \top = \top$, on obtient deux semi-anneaux idempotents complets $(\overline{\mathcal{R}}, \vee, \underline{\divideontimes})$ et $(\overline{\mathcal{R}}, \wedge, \overline{\divideontimes})$. De plus, $(\mathcal{R} \cup \{\perp\}, \vee, \underline{\divideontimes})$ et $(\mathcal{R} \cup \{\top\}, \wedge, \overline{\divideontimes})$ sont des semi-corps idempotents conditionnellement complets. La notion de groupe réticulé conditionnellement complet est en fait équivalente à celle de semi-corps idempotent conditionnellement complet.

Dans le théorème de séparation obtenu dans [J13], le choix de la topologie de $\mathcal{R}$ et $\mathcal{R}^n$ est important. Nous utilisons les notions de topologies de Scott et Lawson introduites en théorie des treillis continus [GHK+80] généralisées au cas des treillis conditionnellement complets dans [J6]. Si $(L, \leq)$ est un treillis conditionnellement complet, on définit la relation d'ordre "*bien inférieur à*" ("way below") $\ll$ par : $a \ll b$ si pour tout sous-ensemble dirigé borné supérieurement D de L, tel que $b \leq \sup D$, il existe $x \in D$ tel que $a \leq x$ (un ensemble dirigé D est tel que tout sous-ensemble fini de D admet un majorant dans D, la relation d'ordre $\ll$ généralise la relation $<$ de $\mathbb{R}$ ou la relation $<$ produit sur $\mathbb{R}^n$). Un treillis conditionnellement complet L est dit *continu* si $x = \sup\{y \in L \mid y \ll x\}$ pour tout $x \in L$. Un sous-ensemble G de L est *montant* (resp. *descendant*) si pour tout $x \in L$ et $y \in G$, l'inégalité $y \leq x$ (resp. $x \leq y$) implique $x \in G$. La *topologie de Scott* sur L est formée des ensembles montants U de L tels que pour tout sous-ensemble dirigé borné supérieurement D de L, $\sup D \in U$ implique $D \cap U \neq \emptyset$. La *topologie inférieure* sur L est générée par les ensembles $\{x \in L \mid a \not\leq x\}$ pour $a \in L$. La *topologie de Lawson* sur L est la topologie générée par l'union des topologies de Scott et inférieure.

Considérons la fonction de couplage $\varphi : \mathcal{R}^n \times \mathcal{R}^n \to \mathcal{R}$ définie par :

$$\varphi(x, w) := \inf_{1 \leq i \leq n} (x_i \divideontimes w_i) \quad \text{pour } x = (x_i) \text{ et } w = (w_i) \in \mathcal{R}^n ,$$

qui n'est autre que le produit scalaire de $\overline{\mathcal{R}}^n$ en tant que semi-module sur le semi-anneau $(\overline{\mathcal{R}}, \wedge, \overline{\divideontimes})$, restreint à $\mathcal{R}^n$. Le résultat suivant généralise les théorèmes de séparation de [RGJ95, MLRS02] (on note ici $\mathbb{1}$ l'élément neutre pour $\divideontimes$) :

Théorème 1.6 ([J13, Theorem 5.1]). *Soit $\mathcal{R} = (\mathcal{R}, \leq, \divideontimes)$ un g.r.c.c. continu et soit G un sous-ensemble de $\mathcal{R}^n$. Les assertions suivantes sont équivalentes :*

1. *G est fermé pour la topologie de Scott de $\mathcal{R}^n$.*
2. *G est descendant et pour tout $x \in \mathcal{R}^n$, l'ensemble*
$$H := \{\lambda \in \mathcal{R} \mid \lambda \divideontimes x \in G\} \tag{1.4}$$
est Scott-fermé dans $\mathcal{R}$.
3. *Pour tout $x \in \mathcal{R}^n \setminus G$, il existe $w \in \mathcal{R}^n$ et $a \in \mathcal{R}$ tels que*
$$a \ll \varphi(x, w),\ a \not\ll \varphi(g, w) \quad \forall g \in G . \tag{1.5}$$

4. Pour tout $x \in \mathcal{R}^n \setminus G$, il existe $w \in \mathcal{R}^n$ tel que

$$\mathbb{1} \ll \varphi(x,w),\, \mathbb{1} \not\ll \varphi(g,w) \quad \forall g \in G\ . \tag{1.6}$$

5. G est descendant et fermé pour la topologie de Lawson de $\mathcal{R}^n$.

6. G est descendant, et pour tout $x \in \mathcal{R}^n$ l'ensemble H défini par (1.4) *est Lawson-fermé dans $\mathcal{R}$.*

Dans le théorème 1.6, $\mathcal{R}$ est un treillis conditionnellement complet continu, ce qui implique en particulier que les topologies de Scott et Lawson passent aux produit : la topologie de Scott de $\mathcal{R}^n$ est aussi la topologie produit sur $\mathcal{R}^n$ de la topologie de Scott sur $\mathcal{R}$. On définit la topologie *Scott duale* comme étant la topologie pour l'ordre opposé $\geq$. La topologie *bi-Scott* est alors celle générée par l'union des topologies Scott et Scott duale. Lorsque $\mathcal{R}$ est un g.r.c.c. continu, cette topologie coïncide avec celle de l'ordre, $(\mathcal{R}, \otimes)$ est un groupe topologique et $(\mathcal{R}, \leq)$ est un treillis topologique [J13, Theorem 4.1]. Dans (5) et (6), on peut remplacer la topologie de Lawson par la topologie bi-Scott.

Dans [J13], nous avons aussi établi un théorème de séparation des ensembles descendants non nécessairement fermés (Proposition 5.1) et des résultats sur les fonctions quasi-convexes généralisées et les conjugaisons de type Lau qui leur sont associées. Ces derniers résultats reposent sur les propriétés des fonctions continues et s.c.i. par rapport aux topologies de Scott et Lawson établies auparavant dans le même article.

1.3 Symétrisation de l'algèbre max-plus

Les travaux présentés ici ont fait l'objet des articles publiés [J1, C3], sous le nom de "Max Plus" donné au groupe de travail sur l'algèbre max-plus de l'INRIA Rocquencourt comprenant au moment de ces travaux Guy Cohen, Stéphane Gaubert, Ramine Nikhoukhah, Jean-Pierre Quadrat et moi-même. Ils ont ensuite été présentés dans [BCOQ92].

Nous nous intéressons ici à la résolution de systèmes d'équations max-plus linéaires de la forme

$$Ax \oplus b = Cx \oplus d\ , \tag{1.7}$$

où l'on cherche $x \in \mathbb{R}^n_{\max}$ et où $A, C \in \mathbb{R}^{p\times n}_{\max}$, $b, d \in \mathbb{R}^p_{\max}$ sont donnés. Ce type de systèmes apparaît dans l'étude de certaines notions de rang max-plus comme le rang défini à partir de l'indépendance de Gondran-Minoux [GM78]. Elle apparaît aussi en théorie des systèmes max-plus, laquelle est utile pour étudier certains systèmes à événements discrets.

Si on était dans l'algèbre usuelle, on pourrait écrire le système (1.7) sous la forme $(A - C)x = (d - b)$ et pour $n = p$, on appliquerait les formules de Cramer. Afin d'utiliser de telles formules dans l'algèbre max-plus, il faut symétriser cette algèbre ou tout au moins donner un sens au signe moins. Or il est bien connu qu'un monoïde idempotent ne peut être symétrisé comme $\mathbb{N}$ ou $\mathbb{R}_+$, puisque tout élément idempotent admettant un opposé est nécessairement nul.

Pourtant la construction classique du symétrisé d'un semi-anneau peut se faire ici jusqu'à un certain point. Celle-ci consiste en effet à considérer l'ensemble des paires $x = (x^+, x^-) \in \mathbb{R}^2_{\max}$, muni de lois correspondantes à l'écriture $x = x^+ \boxminus x^-$. Ainsi, l'addition est l'addition terme à terme, la multiplication vérifie : $x \otimes y = (x^+y^+ \oplus x^-y^-, x^+y^- \oplus x^-y^+)$, et on définit le *signe moins* $\boxminus$ par $\boxminus x = (x^-, x^+)$ (appelé l'*opposé* de x) et $x \boxminus y = x \oplus (\boxminus y)$. On définit aussi la *relation d'équilibre* :

$x \,\Delta\, y \Leftrightarrow x^+ \oplus y^- = x^- \oplus y^+$. Si cette relation était d'équivalence, alors le quotient de $\mathbb{R}^2_{\max}$ par cette relation fournirait le symétrisé de $\mathbb{R}_{\max}$.

La relation Δ n'étant pas transitive, on considère à la place la relation d'équivalence $\mathcal{R}$ définie par :

$$x \,\mathcal{R}\, y \Longleftrightarrow \quad x = y \text{ ou } \left(x^+ \neq x^-,\ y^+ \neq y^-,\ \text{et } x \,\Delta\, y\right) ,$$

qui est plus forte que Δ et qui est compatible avec les lois de $\mathbb{R}^2_{\max}$ ainsi qu'avec le signe moins. En quotientant $\mathbb{R}^2_{\max}$ par cette relation, on obtient un semi-anneau idempotent $\mathbb{S}_{\max}$, appelé *symétrisé* de $\mathbb{R}_{\max}$, qui est l'union de l'ensemble $\mathbb{R}^{\oplus}_{\max}$ des éléments positifs, que l'on peut assimiler à $\mathbb{R}_{\max}$, de l'ensemble $\mathbb{R}^{\boxminus}_{\max}$ des éléments négatifs, qui sont les opposés des éléments positifs, et de l'ensemble $\mathbb{R}^{\bullet}_{\max}$ des éléments équilibrés, qui sont les éléments $x \in \mathbb{S}_{\max}$ tels que $x \,\Delta\, \mathbb{0}$ ou de manière équivalente les éléments de la forme $x^{\bullet} := x \boxminus x$, pour $x \in \mathbb{R}_{\max}$ ou pour $x \in \mathbb{S}_{\max}$. Ces ensembles sont deux à deux d'intersection réduite au singleton $\{\mathbb{0}\}$.

Un élément de $\mathbb{S}_{\max}$ est dit *signé* si il est positif ou négatif. Les éléments signés non nuls sont exactement les éléments inversibles (pour la multiplication) de $\mathbb{S}_{\max}$. Les éléments signés ont l'avantage de vérifier la propriété de transitivité faible suivante [C3, Property 4.3] :

$$(a \,\Delta\, b, b \,\Delta\, c \text{ et } b \text{ signé}) \Rightarrow a \,\Delta\, c \,.$$

En étendant la relation d'équilibre aux vecteurs à coordonnées dans $\mathbb{S}_{\max}$, on obtient facilement que $x \in \mathbb{R}^n_{\max}$ est solution (1.7) si et seulement si il est solution de l'équation $(A \boxminus C)x \,\Delta\, (d \boxminus b)$ [C3, Proposition 4.5].

La construction de $\mathbb{S}_{\max}$ permet en particulier de définir la notion de déterminant dans $\mathbb{R}_{\max}$ et aussi dans $\mathbb{S}_{\max}$, de la manière habituelle. Si on note A^{adj} la transposée de la matrice des cofacteurs de A (que l'on définit de la manière habituelle à partir de la notion de déterminant), on obtient l'identité $AA^{\mathrm{adj}} \,\Delta\, \det A\, \mathbb{I}$ [C3, Theorem 5.1]. Ceci permet de montrer facilement l'implication inverse dans le théorème suivant, à partir de la transitivité faible de la relation d'équilibre.

Théorème 1.7 ([C3, Theorem 6.1])**.** *Soit $A \in \mathbb{S}^{n\times n}_{\max}$ et $b \in \mathbb{S}^n_{\max}$. Alors toute solution signée x de l'équation*

$$Ax \,\Delta\, b \tag{1.8}$$

vérifie nécessairement : $\det A\, x \,\Delta\, A^{\mathrm{adj}} b$*. Inversement, si $A^{\mathrm{adj}} b$ est signé et* $\det A$ *inversible, la solution donnée par "les formules de Cramer",* $(\det A)^{-1} A^{\mathrm{adj}} b$*, est l'unique solution de* (1.8)*.*

Considérons par exemple le système d'équations dans $\mathbb{R}_{\max}$:

$$\begin{cases} \max(x, y-4, 1) = \max(x-1, y+1, 2) \\ \max(x+3, y+2, -5) = \max(y+2, 7) \end{cases} \tag{1.9}$$

qui s'écrit sous forme matricielle dans $\mathbb{R}_{\max}$:

$$\begin{bmatrix} 0 & -4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} 1 \\ -5 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -\infty & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} 2 \\ 7 \end{bmatrix} .$$

Ce système est équivalent à la recherche d'un vecteur (x, y) vérifiant dans $\mathbb{S}_{\max}$:

$$\begin{bmatrix} 0 & \boxminus 1 \\ 3 & 2^{\bullet} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \,\Delta\, \begin{bmatrix} 2 \\ 7 \end{bmatrix} . \tag{1.10}$$

La matrice de ce système ayant un déterminant inversible égal à 4, les formules de Cramer donnent

$$x = \begin{vmatrix} 2 & \boxminus 1 \\ 7 & 2^\bullet \end{vmatrix} \mathbin{/\!\!/} 4 = 8 \mathbin{/\!\!/} 4 = 4, \quad y = \begin{vmatrix} 0 & 2 \\ 3 & 7 \end{vmatrix} \mathbin{/\!\!/} 4 = 7 \mathbin{/\!\!/} 4 = 3$$

qui sont des éléments signés. Le théorème 1.7 montre donc que l'équation (1.10) admet comme unique solution signée $(x, y) = (4, 3)$. Comme les coordonnées de cette solution sont positives, celle-ci est aussi l'unique solution de (1.9) dans $\mathbb{R}_{\max}$.

On définit la *valeur absolue* $|\cdot|$ sur $\mathbb{S}_{\max}$ par quotient de celle définie sur $\mathbb{R}^2_{\max}$ par $|x| = x^+ \mathbin{+\!\!\!+} x^-$. De même on définit la valeur absolue d'un vecteur coordonnée par coordonnée. Le résultat suivant permet de résoudre certains cas dégénérés. La preuve est basée sur un algorithme de type Jacobi (ou de Gauss-Seidel) avec un ordre bien choisi des équations.

Théorème 1.8 ([C3, Theorem 6.2])**.** *Soit $A \in \mathbb{S}^{n\times n}_{\max}$, telle que* $\det A \neq \mathbb{0}$*, et soit $b \in \mathbb{S}^n_{\max}$. Alors il existe une solution signée x à l'équation $Ax \,\nabla\, b$ telle que* $|x| = |\det A|^{-1}|A^{\mathrm{adj}}b|$.

Dans [C3], on montre aussi que dans le cas carré, l'équation $Ax \,\nabla\, \mathbb{0}$ admet une solution signée non nulle si et seulement si $\det A \,\nabla\, \mathbb{0}$.

Le cas des matrices rectangulaires non traité dans ce travail a été étudié par la suite dans le travail de Gaubert [Gau92], qui montre qu'on ne peut pas généraliser les résultats usuels (dans le cas rectangulaire, l'existence d'une solution signée non nulle de l'équation $Ax \,\nabla\, \mathbb{0}$ ne peut être caractérisée à l'aide de déterminants). Dans [GB99], les résultats de symétrisation sont reliés à la théorie de la “résolution par signes” des systèmes linéaires. Voir [DSDM98] pour d'autres travaux sur la symétrisation.

Signalons enfin une interprétation de $\mathbb{S}_{\max}$ en termes d'asymptotiques. Sur $\mathbb{S}_{\max}$ l'addition vérifie : $x \mathbin{+\!\!\!+} y = x$ si $|x| > |y|$ ou $x = y$, $x \mathbin{+\!\!\!+} y = y$ si $|x| < |y|$, et $x \mathbin{+\!\!\!+} y = x^\bullet$ si $x = \boxminus y$. On peut donc voir $\mathbb{S}_{\max}$ comme l'algèbre des développements limités autour de 0^+ de la forme : $f(\varepsilon) = a\varepsilon^{-b} + o(\varepsilon^{-b})$, pour lesquels seul le signe du coefficient $a \in \mathbb{R}^*$ est connu, auxquels on rajoute les fonctions dont on sait seulement qu'elles sont en $O(\varepsilon^{-b})$. Il suffit d'associer b, $\boxminus b$ ou $b^\bullet$ à f selon que $a > 0$, $a < 0$ ou $f(\varepsilon) = O(\varepsilon^{-b})$. On peut aussi remplacer la condition “$f(\varepsilon) = a\varepsilon^{-b} + o(\varepsilon^{-b})$” par les conditions “$f$ est à valeurs réelles de signe constant autour de 0 et $b = \lim_{\varepsilon\to 0^+} \log(|f(\varepsilon)|)/(-\log(\varepsilon))$ existe”.

1.4 Perspectives / Travaux en cours

L'application des résultats de la section 1.1.2 à la caractérisation de taux de grandes déviation est exposée dans la section 3.4.

L'équation fonctionnelle $Bf = g$ apparaît aussi dans la formulation duale du problème de transport de masse de Monge-Kantorovitch, que l'on peut voir comme le calcul du permanent max-plus de l'opérateur B. Dans un travail en cours avec S. Gaubert et V. Kolokoltsov, nous étendons au cas de la dimension infinie “dénombrable” (i.e. lorsque B est un opérateur à noyau avec X et Y dénombrables) les résultats de P. Butkovič (voir par exemple [But00]), sur le lien entre permanent fort d'une matrice max-plus B (ou unicité du problème d'affectation optimale), forme normale de B, et existence d'un élément f ayant une unique pré-image par B (ou existence et unicité du transport optimal). On peut envisager d'étudier dans le même esprit le cas de la dimension infinie “continue” et ses liens avec le problème de Monge-Kantorovitch.

La symétrisation de l'algèbre max-plus, permet de définir une notion de rang de matrice dans $\mathbb{S}_{\max}$ qui généralise le rang de Gondran-Minoux. D'autres symétrisations de l'algèbre max-plus ont été introduites dans la littérature récente en géométrie tropicale. Dans un travail en cours avec S. Gaubert et A. Guterman (Université d'état de Moscou), nous revisitons les différentes notions de rang, symétrisés ou pas, et nous les comparons. Voir aussi [B8] où nous avons déjà établi certaines de ces comparaisons.

L'étude du lien entre symétrisation et asymptotiques mérite aussi d'être repris, par exemple en lien avec les asymptotiques de la valeur propre de Perron étudiées dans la section 5.1.

D'autres développements possibles sur le thème des systèmes linéaires max-plus peuvent être envisagés en lien avec la convexité abstraite, en prolongeant les travaux exposés dans la section 1.2.

Chapitre 2

Théorie spectrale max-plus

Les travaux présentés ici ont été obtenus en collaboration avec Stéphane Gaubert (INRIA Rocquencourt) et Cormac Walsh (INRIA Rocquencourt). Ils ont fait l'objet des articles publiés [B7, C19] et de la pré-publication [S2].

On s'intéresse au problème spectral max-plus $Au = \lambda u$, lorsque $A : \mathbb{R}_{\max}^S \to \overline{\mathbb{R}}_{\max}^S$ est un opérateur max-plus linéaire à noyau fini, sur un espace de dimension infinie, i.e. lorsque S est infini et

$$(Au)_i = \sum_{j \in S} A_{i,j} u_j \quad \forall i \in S \;, \tag{2.1}$$

pour un noyau $A : (i,j) \in S \times S \mapsto A_{i,j} \in \mathbb{R}_{\max}$. On cherche le *vecteur propre* u dans $\mathbb{R}_{\max}^S \setminus \{\mathbb{0}\}$ et la *valeur propre* λ dans $\mathbb{R}_{\max}$. En notations usuelles, le problème spectral s'écrit :

$$\sup_{j \in S} (A_{i,j} + u_j) = \lambda + u_i \quad \forall i \in S \;. \tag{2.2}$$

Le problème spectral (2.2) intervient en contrôle ergodique, où l'on cherche le maximum du gain moyen par unité de temps en temps long, dans l'étude de systèmes à événements discrets [BCOQ92, HOvdW05], en mécanique statistique [CG86], dans l'étude de systèmes à retard [MPN03, MPN02], et dans les perturbations de valeurs propres (cf. le chapitre 5).

Le cas où S est fini a été traité par de nombreux auteurs dans la littérature [CG79, Rom67, Vor67, GM77, CDQV83]. On a alors une caractérisation de la valeur propre dans le cas irréductible comme le poids moyen maximal d'un circuit. On a aussi un résultat de représentation de l'espace propre, et une description précise du comportement asymptotique des itérées de A, à l'aide d'un certain graphe, dit graphe critique et de sa cyclicité. Des résultats existent également lorsque S est compact et que le noyau vérifie certaines propriétés de régularité ou certaines conditions géométriques. Nous considérons ici le cas d'un ensemble S non compact, cas qui a été très peu étudié dans la littérature jusqu'ici. Nous nous intéressons aussi au cas de semi-groupes à temps continu d'opérateurs max-plus linéaires à noyau.

2.1 Théorie spectrale max-plus dénombrable

Dans [B7], nous avons établi quelques résultats généraux de théorie spectrale en dimension infinie. Nous avons ensuite montré, dans le cas où S est dénombrable, et sous des hypothèses particulières de croissance à l'infini, que l'on appelle tension, un résultat de représentation de l'espace propre et un résultat de cyclicité qui généralisent ceux établis dans la littérature pour le cas où S est fini. Nous présentons ici ces deux résultats.

Étant donné un noyau, aussi considéré comme une matrice, $A \in \mathbb{R}_{\max}^{S\times S}$, on définit le graphe de A, $\mathsf{G}(A)$, de la manière habituelle : l'ensemble des nœuds est S, et il existe un arc du nœud i au nœud j si $A_{ij} \neq \mathbb{0}$. La matrice A est dite *irréductible* si son graphe est fortement connexe. Le *poids* d'un arc (i,j) dans ce graphe est donné par $A_{i,j}$. Le *poids* d'un chemin $w = (i_0, i_1, , \ldots, i_{k-1}, i_k)$ est donné par $|w|_A := A_{i_0,i_1} \cdots A_{i_{k-1},i_k}$, qui est donc la somme des poids de ses arcs dans algèbre usuelle, et sa *longueur* est égale à $|w| := k$. Le chemin w est un *circuit* si $i_k = i_0$. Son *poids moyen* est égal à $|w|_A/|w|$. On note $\rho_{\max}(A)$ le maximum des poids moyens des circuits de $\mathsf{G}(A)$. Le *graphe critique* $\mathsf{GC}(A)$, est le sous-graphe de $\mathsf{G}(A)$ obtenu en prenant l'union de tous les circuits de poids moyen maximal, appelés *circuits critiques*. On note $N^c(A)$ l'ensemble des *nœuds critiques* de A, qui sont les nœuds de $\mathsf{GC}(A)$.

Si A a un vecteur propre $u \in \mathbb{R}^S$ (i.e. qui ne prend jamais la valeur $\mathbb{0}$) associé à la valeur propre λ, alors $\lambda \geq \rho_{\max}(A)$. Par contre on n'a pas en général l'égalité, comme c'est le cas lorsque S est fini, même dans le cas où A est irréductible. On peut cependant garantir l'unicité de λ, en imposant une condition de tension que nous décrivons ci-après.

Étant donné un vecteur $v = (v_j)_{j\in S} \in \overline{\mathbb{R}}_{\max}^{S}$, on dit qu'il est *tendu* si il appartient à $\mathbb{R}_{\max}^S$, et si pour tout $\beta \in \mathbb{R}$, l'*ensemble de sur-niveau* $J_\beta := \{j \in S \mid v_j \geq \beta\}$ est fini. Le mot "tendu" réfère à la tension de la mesure idempotente de densité v, cf. les sections 3.2 et 3.4. Si $B \in \overline{\mathbb{R}}_{\max}^{S\times S}$ est une matrice, on dit que v est *B-tendu* si, pour tout $i \in S$, le vecteur $(B_{ij}v_j)_{j\in S}$ est tendu. On dit alors qu'une matrice A *a la propriété* (T) si tous les vecteurs colonnes de $A^\star := \mathbb{I} + A + A^2 + \cdots$ sont $A^\star$-tendus. Cette condition est équivalente à la condition que pour tous $i,j \in S$, $A^\star_{ik}A^\star_{kj}$ tends vers $\mathbb{0}$ lorsque k va à l'infini, elle implique que les trajectoires optimales restent à distance finie, et dans le cas irréductible, elle implique que S est dénombrable.

La notion suivante est analogue à la récurrence des chaînes de Markov. Considérons $\tilde{A} := \rho_{\max}(A)^{-1}A$, et pour toute matrice B, notons $B^+ = BB^\star = B + B^2 + \cdots$. On dit que $i \in S$ est *récurrent* si $\tilde{A}^+_{i,i} = \mathbb{1}$, et on appelle *classes de récurrence* les classes d'équivalence pour la relation $\mathcal{R}$ définie par $i\mathcal{R}j$ si $\tilde{A}^+_{ij}\tilde{A}^+_{ji} = \mathbb{1}$. Dans le cas où S est fini, les nœuds récurrents sont exactement les nœuds critiques et les classes de récurrence correspondent aux composantes fortement connexes du graphe critique. Dans le cas infini, ceci est encore le cas, si $\tilde{A}$ a la propriété (T) [B7, Theorem 4.9].

Le théorème suivant est un résultat de représentation des vecteurs propres "tendus" généralisant le théorème de représentation établi dans la littérature dans le cas où S est fini.

Théorème 2.1 ([B7, Theorem 6.1])**.** *Soit $u \in \mathbb{R}_{\max}^S \setminus \{\mathbb{0}\}$ un vecteur propre de A associé à une valeur propre $\lambda \in \mathbb{R}$ telle que $\lambda \geq \rho_{\max}(A)$. Supposons que u est $(\lambda^{-1}A)^\star$-tendu. Alors $\lambda = \rho_{\max}(A)$, l'ensemble $N^c(A)$ des nœuds critiques de A est non vide et*

$$u = \sum_{j\in N^c(A)} \tilde{A}^\star_{\cdot j} u_j \; . \tag{2.3}$$

Si A est irréductible, les hypothèses du théorème 2.1 impliquent que $\tilde{A}$ a la propriété (T). On montre aussi que, sous cette dernière condition, une famille de vecteurs colonnes $\tilde{A}^\star_{\cdot j}$ est une famille génératrice minimale du semi-module $\mathscr{T}$ des vecteurs $\tilde{A}^\star$-tendus u tels que $Au = \rho_{\max}(A)u$, si l'on choisit exactement un élément j dans chaque classe critique de A [B7, Theorem 6.5]. Ceci implique en particulier que les vecteurs colonnes $\tilde{A}^\star_{\cdot j}$ pour j critique sont des générateurs extrémaux de $\mathscr{T}$. Ici et plus loin, on dit qu'un vecteur u d'un semi-module $\mathscr{T}$ est un *générateur extrémal*, si $u = v \oplus w$ et $v, w \in \mathscr{T}$ impliquent $u = v$ ou $u = w$. Cette notion d'extrémalité peut être vue comme un analogue max-plus de la notion de direction extrémale d'un cône convexe.

On note $\sigma(A)$ la cyclicité du graphe critique de A. Rappelons que si un graphe est fortement connexe, sa cyclicité est le p.g.c.d. des longueurs des circuits de ce graphe, et sinon c'est le p.p.c.m. des cyclicités des composantes fortement connexes de ce graphe. Le théorème suivant généralise le théorème de cyclicité établi dans [CDQV83] dans le cas où S est fini.

Théorème 2.2 ([B7, Theorem 7.4])**.** *On suppose que A est irréductible, que $\tilde{A}$ a la propriété* (T)*, et que l'ensemble $N^c(A)$ des nœuds critiques de A est non vide. Alors pour tous $i, j \in S$, il existe $\sigma_{ij} \in \mathbb{N} \setminus \{0\}$ et $n_{ij} \in \mathbb{N}$ tels que*

$$A_{ij}^{n+\sigma_{ij}} = \rho_{\max}(A)^{\sigma_{ij}} A_{ij}^n \quad \text{pour } n \geq n_{ij} \;, \tag{2.4}$$

et on a la formule explicite (de type projecteur spectral)

$$\tilde{A}_{ij}^n = \sum_{k \in N^c(A)} (\tilde{A}^q (\tilde{A}^{\sigma_{ij}})^\star_{ik} (\tilde{A}^{q'} (\tilde{A}^{\sigma_{ij}})^\star)_{kj} \;, \quad \text{pour } n \geq n_{ij}, \tag{2.5}$$

où q, q' sont des nombres arbitraires dans $\{0, \ldots, \sigma_{ij} - 1\}$ tels que $q + q' \equiv n\,[\sigma_{ij}]$. De plus, si $\sigma(A)$ est fini, l'entier σ_{ij} peut être choisi divisant $\sigma(A)$.

On obtient aussi une estimation précise des temps de couplage n_{ij}.

2.2 La frontière de Martin max-plus

On veut ici trouver, pour toute valeur propre $\lambda \in \mathbb{R}$ de A, une représentation de l'ensemble de tous les vecteurs propres associés et non pas seulement des vecteurs propres tendus. Sans perte de généralité, on peut considérer le cas où $\lambda = \mathbb{1}$. On se ramène donc à l'étude des points fixes, que l'on appelle fonctions harmoniques par analogie avec la théorie (classique ou probabiliste) du potentiel. Dans [S2], nous avons construit un analogue à la frontière de Martin probabiliste, et obtenu un théorème de représentation des fonctions harmoniques à partir de cette frontière de Martin qui est analogue celui établi pour les fonctions harmoniques usuelles. Nous présentons ici ce résultat.

Pour cela, nous avons besoin de supposer qu'il existe un vecteur ligne $\pi \in \mathbb{R}^S$ qui soit sur-harmonique, i.e. tel que $\pi \geq \pi A$, ce qui implique en particulier que $\rho_{\max}(A) \leq \mathbb{1}$. On note $\mathscr{H}$ le semi-module des fonctions harmoniques u de A qui sont π-intégrables, i.e. qui vérifient $\pi u < +\infty$. Si A est irréductible, on peut choisir $\pi := A^\star_{b\cdot}$ pour un certain point de base $b \in S$. Dans ce cas toute fonction harmonique est nécessairement π-intégrable, et donc $\mathscr{H}$ est l'ensemble de toutes les fonctions harmoniques par rapport à A.

On définit le *noyau de Martin* K par rapport à π par :

$$K_{ij} := A^\star_{ij} (\pi_j)^{-1} \quad \text{pour tous } i, j \in S \;.$$

On montre facilement que les colonnes $K_{\cdot j}$ sont toutes bornées supérieurement par un même vecteur de $\mathbb{R}^S$. Ceci implique que l'ensemble $\mathscr{K}$ des colonnes $K_{\cdot j}$ est relativement compact pour la topologie produit (qui est aussi la topologie de la convergence simple). L'*espace de Martin* est par définition l'adhérence de $\mathscr{K}$. C'est un compact, et on note $\mathscr{B} := \mathscr{M} \setminus \mathscr{K}$ la *frontière de Martin.* La frontière de Martin max-plus généralise dans une certaine mesure la frontière d'un espace métrique construite à partir des horofonctions (fonctions de Busemann généralisées), introduite par Gromov, et dont l'analogie avec la frontière de Martin probabiliste a été notée par Ballman. Par contre, elle n'est pas en général fermée, et l'application $j \in S \mapsto K_{\cdot j} \in \mathscr{M}$, qui est une injection par exemple lorsque $\rho_{\max}(A) < \mathbb{1}$ ou lorsque A n'a pas de nœuds récurrents, n'est pas forcément un plongement [S2, Example 10.6].

Si $u \in \mathbb{R}_{\max}^S$ est π-intégrable, on définit la fonction $\mu_u : \mathscr{M} \to \mathbb{R}_{\max}$ par

$$\mu_u(w) := \limsup_{K_{\cdot j} \to w} \pi_j u_j := \inf_{W \ni w} \sup_{K_{\cdot j} \in W} \pi_j u_j \quad \text{pour } w \in \mathscr{M} \ ,$$

où l'infimum est pris sur tous les voisinages ouverts W de w dans $\mathscr{M}$. La fonction μ_u est une fonction s.c.s. sur $\mathscr{M}$ à valeurs dans $\mathbb{R}_{\max}$.

Afin d'obtenir un théorème de représentation "minimal" similaire au théorème 2.1, on introduit sur $\mathscr{M}$ un noyau $H^\flat$ qui prolonge A^+ de la façon suivante :

$$H^\flat(w', w) := \limsup_{K_{\cdot i} \to w'} \liminf_{K_{\cdot j} \to w} \pi_i A_{ij}^+ (\pi_j)^{-1} \ .$$

Ce noyau vérifie $H^\flat(w', w) \leq \mathbb{1}$ pour tous $w, w' \in \mathscr{M}$ et $H^\flat(K_{\cdot i}, K_{\cdot j}) = \pi_i A_{ij}^+ (\pi_j)^{-1}$. On définit alors l'*espace de Martin minimal* comme l'ensemble $\mathscr{M}^m := \{w \in \mathscr{M} \mid H^\flat(w, w) = \mathbb{1}\}$. Notons que l'on peut aussi définir sur $\mathscr{M}$ un noyau H qui prolonge $A^\star$ par : $H(w', w) := \mu_w(w')$ Alors, $H(w, w) = H^\flat(w, w)$ si $w \in \mathscr{B}$ et $H(w, w) = \mathbb{1}$ sinon. De plus, $\mathscr{M}^m$ est l'ensemble des éléments w de la fontière de Martin $\mathscr{B}$ qui vérifient $H(w, w) = \mathbb{1}$ auquel on rajoute éventuellement les colonnes $K_{\cdot j}$ pour j récurrent, lorsque $\rho_{\max}(A) = \mathbb{1}$.

Le théorème de représentation suivant est analogue au théorème de représentation de fonctions harmoniques classiques, la fonction μ_u jouant le rôle de la mesure spectrale :

Théorème 2.3 ([S2, Theorem 8.1])**.** *Tout élément $u \in \mathscr{H}$ peut s'écrire*

$$u = \sum_{w \in \mathscr{M}^m} \nu(w) w \ , \tag{2.6}$$

avec $\nu : \mathscr{M}^m \to \mathbb{R}_{\max}$, et alors

$$\sup_{w \in \mathscr{M}^m} \nu(w) < +\infty \ . \tag{2.7}$$

Inversement, toute fonction $\nu : \mathscr{M}^m \to \mathbb{R}_{\max}$ vérifiant (2.7) *définit par* (2.6) *un élément u de $\mathscr{H}$. De plus, pour tout $u \in \mathscr{H}$, μ_u est la fonction maximale ν vérifiant* (2.6)*.*

Les générateurs extrémaux normalisés u (i.e. vérifiant $\pi u = \mathbb{1}$) de $\mathscr{H}$ sont exactement les éléments de $\mathscr{M}^m$ [S2, Theorem 8.3]. De plus les éléments de l'espace de Martin minimal peuvent être caractérisés comme les limites de certaines (quasi) géodésiques [S2, Corollary 7.7 et Proposition 7.8]. Ces résultats ne semblent pas avoir été établis dans le contexte des frontières d'espaces métriques.

2.3 La frontière de Martin max-plus de semi-groupes à temps continu

Nous considérons ici un semi-groupe à temps continu $(A^t)_{t \geq 0}$ d'opérateurs max-plus linéaires à noyau sur $\mathbb{R}_{\max}^S$. On appelle *vecteur propre* de A^t associé à la valeur propre $\lambda \in \mathbb{R}$, un élément

$u \in \mathbb{R}_{\max}^S \setminus \{\mathbb{0}\}$ tel que $A^t u = \lambda^t u$ en notations max-plus, i.e. tel que

$$\sup_{y \in S} \big(A^t(x,y) + u(y) \big) = \lambda t + u(x) \quad \forall x \in S \ , \tag{2.8}$$

en notations usuelles, où on note ici $A^t(x,y)$ le noyau de l'opérateur A^t aux points $x, y \in S$.

Tout semi-groupe de Lax-Oleinik associé à un problème de contrôle optimal déterministe est un cas particulier de semi-groupe à temps continu d'opérateurs max-plus linéaires à noyau. Dans le cas où S est une variété et où le lagrangien vérifie des propriétés de régularité et de convexité forte, l'étude des vecteurs propres fait l'objet de la théorie KAM faible de Fathi [Fat97, Fat03], ceux-ci étant appelés solutions KAM faibles. Un théorème de représentation similaire au théorème 2.3 a été établi dans le cas où S est compact. Dans le cas non compact, Contreras [Con01] a défini un analogue à l'espace de Martin minimal au moyen de fonctions de Buseman, et a établi un résultat de représentation au moyen de cet espace. Les propriétés d'extremalité ne semblent pas avoir été considérées dans ces travaux.

Dans [C19], nous avons présenté un résultat de représentation similaire au théorème 2.3 pour les fonctions harmoniques d'un semi-groupe $(A^t)_{t\geq 0}$, sous une hypothèse très faible, qui inclue en particulier tout semi-groupe de Lax-Oleinik sans hypothèse de régularité du lagrangien. Bien sûr la régularité serait nécessaire à l'existence de géodésiques, mais nous n'avons besoin ici que de quasi-géodésiques. Les preuves complètes font l'objet d'un article en préparation [P7].

Il faut noter que même si les théorèmes de représentation pour les opérateurs max-plus linéaires et pour les semi-groupes à temps continu d'opérateurs max-plus linéaires sont similaires, on ne peut déduire l'un à partir de l'autre, sauf dans des cas très particuliers comme le cas d'un lagrangien invariant en espace (voir la section 12 de [S2]).

Nous présentons brièvement le cadre des résultats de [C19]. Nous considérons comme dans la section précédente le cas de la valeur propre $\lambda = \mathbb{1}$, i.e. 0, et utilisons les notations usuelles. On définit alors les *fonctions harmoniques* comme les solutions $u \in \mathbb{R}_{\max}^S$ de (2.8) pour $\lambda = 0$.

L'étoile de Kleene est maintenant remplacée par le noyau potentiel $A^\star : (x,y) \in S \times S \mapsto \sup_{t\geq 0} A^t(x,y)$. On suppose que celui-ci ne prend que des valeurs finies, ce qui est une condition d'irréductibilité. On peut alors considérer un noyau de Martin avec point de base b : $K(x,y) = A^\star(x,y) - A^\star(b,y)$. L'*espace de Martin* est défini, comme dans la section précédente, comme l'adhérence de $\mathscr{K} := \{K(\cdot,y) \mid y \in S\}$. Comme il est difficile de définir l'analogue du noyau $H^\flat$, on se contente de définir H comme dans la section précédente, et l'*espace de Martin minimal* $\mathscr{M}^m$ comme l'ensemble des fonctions w de $\mathscr{M}$ qui sont harmoniques et qui vérifient $H(w,w) = 0$.

On appelle *chemin* de S une application d'un intervalle de $\mathbb{R}_+$ dans S. On dit qu'une fonction I de l'espace des chemins de S dans $\mathbb{R}_{\max}$ est *additive* si $I(\gamma\gamma') = I(\gamma) + I(\gamma')$ pour tous chemins γ et γ' de S tels que γ finit au point où le chemin γ' commence, où $\gamma\gamma'$ désigne la concaténation des chemins γ et γ'. L'unique hypothèse à faire, en plus de la finitude de $A^\star$, est l'existence d'une fonction (de gain) I additive de l'espace des chemins de S dans $\mathbb{R}_{\max}$ telle que, pour tous $t \in \mathbb{R}_+$ et $x, y \in S$, $A^t(x,y) = \sup_\gamma\{I(\gamma)\}$, où le supremum est pris sur tous les chemins $\gamma : [0,t] \to S$ allant de x à y. Cette hypothèse est vérifiée en particulier si A^t est le semi-groupe de Lax–Oleinik d'un problème de contrôle optimal déterministe à temps continu, mais on peut aussi considérer des fonctions de gain I dont le support contient des chemins avec sauts. Sous ces hypothèses, on obtient un théorème de représentation analogue au théorème 2.3.

2.4 Perspectives / Travaux en cours

Dans le prolongement des résultats de la section 2.2, la caractérisation de la frontière de Martin max-plus de certains espaces normés, et l'existence d'une mesure spectrale minimale ont été établis par Cormac Walsh seul [Wal07, Wal05]. D'autres travaux sont en germe en collaboration avec Stéphane Gaubert et Roger Nussbaum sur l'espace propre d'opérateurs max-plus linéaires sur un espace compact mais à noyau non régulier. On cherche dans ce cas l'ensemble des fonctions propres continues.

Dans le prolongement des travaux présentés dans la section 2.3, il faudrait obtenir l'équivalence entre les vecteurs propres de semi-groupes max-plus linéaires et les solutions de viscosité d'équations d'Hamilton-Jacobi stationnaires, sous des hypothèses faibles de régularité, afin de généraliser les résultats de Fathi et Siconolfi [FS04]. Le cas de problèmes avec contrôles singuliers, ou plus généralement avec fonctions de gain dont le support contient des chemins avec sauts, devrait conduire à des inéquations variationnelles ou quasi-variationnelles, cas qui n'a pas été étudié dans les travaux sur les solutions KAM faible.

Enfin, afin d'obtenir une représentation complète de l'espace des vecteurs propres non linéaires d'un opérateur monotone additivement homogène convexe correspondant à un problème de contrôle stochastique à temps discret ou continu (cf. la section 6.1) sur un espace infini, non compact, il faudrait introduire une compactification qui serait à la fois analogue à l'espace de Martin probabiliste et à l'espace de Martin max-plus et qui engloberait d'une certaine manière les deux compactifications. On peut se demander jusqu'à quel point les résultats de ce chapitre peuvent être généralisés dans ce cadre.

Chapitre 3

Probabilités idempotentes et grandes déviations

Si l'on considère la théorie de la mesure dans laquelle le semi-corps $(\mathbb{R}_+, +, \times)$ est remplacé par un semi-anneau idempotent, on obtient la notion de mesure idempotente introduite par Maslov [Mas73]. Sur l'algèbre min-plus, la mesure d'un ensemble correspond au minimum d'une fonction sur cet ensemble. La théorie des probabilités correspond alors à la théorie de l'optimisation : les variables aléatoires opèrent comme des changements de variables ou des paramétrisations de problèmes d'optimisation, la propriété de Markov correspond au principe de la programmation dynamique (Bellman), la convergence faible à une convergence de type épigraphe.... Les théorèmes limites associés aux distributions stables (loi des grands nombres et théorème de la limite centrale) fournissent des résultats de convergence en contrôle optimal. Au delà de l'analogie, la transformée de Cramér introduite en théorie des grandes déviations, établit un morphisme entre probabilités et optimisation. Finalement, on peut englober probabilités et optimisation dans une même théorie, dont les grandes déviations serait le liant.

Nous présentons ici quelques notions et résultats relatifs à ces idées.

3.1 Densité de probabilités idempotentes

Les travaux présentés ici ont fait l'objet de l'article publié [J6].

On considère ici la notion de mesure idempotente introduite par Maslov [Mas73] sur un semi-anneau idempotent, et on se pose la question de l'existence d'une densité pour cette mesure. Il s'agit d'obtenir un théorème de représentation analogue au théorème 1.1, lorsqu'on considère des mesures ou fonctions dénombrablement additives (alors que dans le théorème 1.1, on considérait des fonctions infiniment additives).

On dit qu'un ensemble $\mathscr{A}$ de parties d'un ensemble donné Ω est une *semi-algèbre de Boole* (resp. *semi-σ-algèbre*) si Ω et $\emptyset$ sont dans $\mathscr{A}$ et si $\mathscr{A}$ est stable par les opérations union finie (resp. dénombrable) et intersection finie. L'ensemble des ouverts d'un espace topologique est une semi-σ-algèbre.

On considère un semi-anneau idempotent $(\mathbb{D}, \oplus, \otimes)$ conditionnellement complet (cf. la section 1.2). Il existe alors un semi-anneau idempotent complet $\overline{\mathbb{D}} = \mathbb{D} \cup \{\top\}$ dont $\mathbb{D}$ est un sous-semi-anneau (voir par exemple la section 1.2). Une $\mathbb{D}$-*mesure idempotente* sur une semi-algèbre de Boole $\mathscr{A}$ est une application $\mathbb{K} : \mathscr{A} \to \overline{\mathbb{D}}$ telle que

1. $\mathbb{K}(\emptyset) = \mathbb{0}$,
2. $\mathbb{K}(A \cup B) = \mathbb{K}(A) \oplus \mathbb{K}(B)$ pour tous $A,\ B \in \mathscr{A}$,
3. $\lim_{n\to+\infty} \mathbb{K}(A_n) = \mathbb{K}(A)$ si la suite $A_n \in \mathscr{A}$ est croissante et telle que $A = \cup_{n\in\mathbb{N}} A_n \in \mathscr{A}$.

La mesure $\mathbb{K}$ est dite *finie* si $\mathbb{K}(\Omega) \in \mathbb{D}$.

Si c est une fonction de Ω dans $\overline{\mathbb{D}}$, l'application $\mathbb{K} : \mathcal{P}(\Omega) \to \overline{\mathbb{D}}$ définie par

$$\mathbb{K}(A) = \sup\{c(\omega) \mid \omega \in A\} \tag{3.1}$$

est une $\mathbb{D}$-mesure idempotente sur $\mathcal{P}(\Omega)$. On dit qu'une mesure idempotente $\mathbb{K}$ sur $\mathscr{A}$ a pour *densité* c si elle vérifie (3.1) pour tout $A \in \mathscr{A}$.

Si $\mathscr{A}$ est une semi-algèbre de Boole de sous-ensembles de Ω, alors toute mesure $\mathbb{K}$ sur $\mathscr{A}$ admet une unique extension $\mathbb{K}^+$ à la semi-σ-algèbre $\mathcal{G}$ engendrée par $\mathscr{A}$. Elle admet aussi une extension maximale $\mathbb{K}^*$ à $\mathscr{P}(\Omega)$ [Mas73] par

$$\mathbb{K}^*(A) = \inf_{G\in\mathcal{G}, G\supset A} \mathbb{K}^+(G).$$

On déduit de ce résultat que si $\mathbb{K}$ a une densité sur $\mathscr{A}$, alors $c^* : \omega \mapsto \mathbb{K}^*(\{\omega\})$ est la densité maximale de $\mathbb{K}$ sur $\mathscr{A}$ [J6, Proposition 3.5]. Si de plus $\mathscr{A}$ est une topologie sur Ω, alors c^* est aussi l'unique densité de $\mathbb{K}^*$ sur $\mathscr{P}(\Omega)$ [J6, Proposition 3.7].

Le résultat principal de [J6] est le suivant :

Théorème 3.1 ([J6, Theorem 3.9])**.** *Soit $\mathscr{A}$ une semi-algèbre de Boole de sous-ensembles de Ω vérifiant la propriété :*

> *pour tout $A \in \mathscr{A}$ et tout recouvrement $A \subset \bigcup_{i\in I} A_i$ par des éléments de $\mathscr{A}$, il existe un sous-recouvrement dénombrable de A : $A \subset \bigcup_{i\in J} A_i$ ($J \subset I$ et J dénombrable).*

Alors, toute $\mathbb{D}$-mesure idempotente $\mathbb{K}$ sur $\mathscr{A}$, telle que $[\mathbb{0}, \mathbb{K}(\Omega)]$ est un treillis (complet) dualement continu (i.e. continu pour l'ordre opposé à l'ordre naturel de $\mathbb{D}$), admet c^* comme densité.

Ce résultat admet plusieurs spécialisations, en particulier lorsque $\mathscr{A}$ est une topologie sur Ω. Les hypothèses du théorème sont par exemple vérifiées pour toute mesure sur $\mathbb{R}_{\max}$ (finie ou non), si $\mathscr{A}$ est une topologie sur Ω admettant une base dénombrable d'ouverts. On peut aussi considérer l'ensemble $\mathscr{A}$ des ouverts qui sont unions dénombrables de fermés, lorsque Ω est un espace topologique qui est union dénombrable de compacts [J6, Corollary 3.22].

Dans [J6], on généralise la construction de l'intégrale idempotente de Maslov par rapport à une mesure idempotente au cas d'un semi-anneau $\mathbb{D}$ qui est un treillis conditionnellement complet continu ou dualement continu, tel que la multiplication distribue par rapport aux suprema bornés supérieurement. On donne aussi des conditions suffisantes, pour que cette intégrale s'exprime à l'aide d'une densité. Si $\mathbb{K}$ est finie, il existe une unique forme $\mathbb{D}$-linéaire $\mathbb{V}$ prolongeant $\mathbb{K}$ sur l'espace $\mathcal{I}(\Omega, \mathscr{A})$ des combinaisons linéaires dénombrables de fonctions indicatrices $\mathbb{1}_A$ d'ensembles $A \in \mathscr{A}$ (définies par $\mathbb{1}_A(y) = \mathbb{1}$ si $y \in A$ et $\mathbb{1}_A(y) = \mathbb{0}$ sinon) et qui préserve les limites croissantes. L'espace $\mathcal{I}(\Omega, \mathscr{A})$ coïncide dans certains cas avec l'ensemble des fonctions semi-mesurables, qui sont les fonctions $f : \Omega \to \mathbb{D}$ telles que $\Omega_f(a) := \{\omega \in \Omega,\ a \ll f(\omega)\} \in \mathcal{G}$ pour tout $a \in \mathbb{D}$, où $\mathcal{G}$ est la semi-σ-algèbre engendrée par $\mathscr{A}$. Si $\mathbb{D}$ est continu, alors $\mathbb{V}$ est donnée par : $\mathbb{V}(f) = \sum_{a\in\mathbb{D}} a \otimes \mathbb{K}^+(\Omega_f(a))$. Si $\mathbb{D}$ est dualement continu et $\mathbb{K}^+$ admet une densité c^*, alors $\mathbb{V}$ admet c^* comme densité dans le sens où : $\mathbb{V}(f) = \sum_{\omega\in\Omega} f(\omega) \otimes c^*(\omega)$.

On établit aussi un théorème de représentation de Riesz. Ces résultats permettent de relier l'existence d'une densité d'une mesure ou intégrale idempotente au résultat de représentation établi par Kolokoltsov et Maslov [KM87] pour les applications max-plus linéaires continues sur l'espace des fonctions continues sur Ω.

Finalement, le théorème 3.1 permet de faire le lien entre la notion de mesure idempotente et celle de capacité telle qu'elle est définie par O'Brien et Vervaat [OV91]. Il permet aussi d'établir une caractérisation d'existence d'une fonction de taux de grandes déviations [J6, Section 5].

3.2 Processus de Bellman

Les travaux présentés ici ont été obtenus pour la plupart en collaboration avec Jean-Pierre Quadrat (INRIA Rocquencourt) et Michel Viot (CNRS, Grenoble). Ils ont fait l'objet des articles publiés [B3, C9, B5] et du Rapport de Recherche INRIA [R1]. Voir aussi les articles publiés [C5, J4] sous le nom collectif Max Plus, qui désignait au moment du travail Guy Cohen, Stéphane Gaubert, Jean-Pierre Quadrat, Michel Viot, et moi-même, rejoint par Michael Mc Gettrick pour le deuxième article.

Les notions de mesures et intégrales idempotentes de Maslov, combinées aux résultats d'existence de densité de la section précédente, nous ont permis de construire un formalisme pour l'optimisation analogue à celui des probabilités, dont nous donnons ici une description rapide.

Variables de décision

On appelle *probabilité idempotente*, une mesure idempotente $\mathbb{K}$ sur un semi-anneau $\mathbb{D}$, telle que $\mathbb{K}(\Omega) = \mathbb{1}$. Un *espace de décision* $(\Omega, \mathscr{A}, \mathbb{K})$ est composé d'un ensemble non vide Ω, d'une semi-σ-algèbre $\mathscr{A}$ de sous-ensembles de Ω et d'une $\mathbb{D}$-probabilité idempotente $\mathbb{K}$. On supposera ici, pour simplifier, que $\mathbb{D}$ est le semi-anneau min-plus $\mathbb{R}_{\min}$, et que $\mathscr{A}$ est une topologie sur Ω admettant une base dénombrable d'ouverts. La mesure $\mathbb{K}$ est aussi appelée *mesure de coût*. Elle admet une extension maximale $\mathbb{K}^*$ pour l'ordre naturel de $\mathbb{R}_{\min}$, et cette extension admet une unique densité c^*. Comme l'ordre naturel de $\mathbb{R}_{\min}$ est l'ordre opposé à l'ordre de $\overline{\mathbb{R}}$, $\mathbb{K}^*$ est en fait une extension minimale de $\mathbb{K}$. De plus, la densité c^* est une fonction s.c.i. de Ω dans $\mathbb{R}_{\min}$ dont l'infimum est égal à 0. Une telle fonction sera appelée une *densité de coût*. Par analogie entre les probabilités usuelles et les probabilités idempotentes, on peut définir les notions de *coût conditionnel*, et d'*indépendance* d'événements.

L'analogue d'une variable aléatoire sur $(\Omega, \mathscr{A}, \mathbb{K})$ à valeurs dans un espace topologique $(E, \mathscr{B})$, appelé *variable de décision*, est simplement une application $X : \Omega \to E$. Elle induit une mesure de coût sur $(E, \mathscr{B})$ définie par $\mathbb{K}_X(A) = \mathbb{K}^*(X^{-1}(A))$ pour $A \in \mathscr{B}$, dont la densité "maximale" sera notée c_X et appelée la *densité de coût* de X. Si $E = \mathbb{R}_{\min}$, X est appelée *variable de coût*, et son intégrale $\mathbb{V}(X)$ par rapport à $\mathbb{K}$ est appelée *valeur de X*. Ces notions permettent de définir l'égalité presque sûre pour des variables de décision X et Y à valeurs dans un même espace E, par $\mathbb{K}^*(X \neq Y) = \mathbb{0}$, ainsi que la *convergence presque sûre*, la *converge faible*, et, lorsque E est un espace métrique, la *convergence en probabilité*, appelée ici *convergence en coût*, de variables de décisions. Par exemple X_n converge faiblement vers X si $\lim_{n\to\infty} \mathbb{V}(f(X_n)) = \mathbb{V}(f(X))$ pour toute fonction $f : E \to \mathbb{R}_{\min}$ continue bornée (inférieurement). On définit de manière similaire la convergence faible de mesures de coût ou de densités de coût.

La valeur d'une variable de coût constitue un premier analogue à la notion d'espérance, que l'on peut aussi obtenir comme la limite d'espérances de certaines variables aléatoires réelles positives.

Par exemple, si P_n est une suite de probabilités vérifiant un principe des grandes déviations de facteurs α_n et de taux c^*, le principe de contraction de Varadhan implique que si X est continue bornée inférieurement alors $\mathbb{V}(X) = \lim_{n\to\infty} -\alpha_n \log \left(\mathbb{E}_{P_n}(\exp(\frac{-X}{\alpha_n}))\right)$ (voir la section 3.4 pour un aperçu et une généralisation des notions utilisées ici).

Un deuxième analogue de la notion d'espérance est la notion d'*optimum* d'une variable de décision X à valeur dans un e.v.t.l.c.s. E quelconque défini par :

$$\mathbb{O}(X) := \operatorname*{Argmin}_{x\in E} \operatorname{conv}(c_X)(x)$$

lorsqu'il existe. Ici conv désigne l'enveloppe convexe s.c.i., et Argmin désigne l'ensemble des points où le minimum de la fonction est atteint. Supposons que E et E' sont en dualité. Comme l'algèbre min-plus s'obtient comme limite, lorsque α tend vers 0^+, des transformées de l'algèbre des nombres réels positifs par les applications $x \mapsto -\alpha \log x$, l'analogue à la transformée de Laplace d'une variable aléatoire X à valeurs dans E, $\mathcal{L}(X) : E' \to \mathbb{R}_+ \cup \{+\infty\}$, $\theta \mapsto \mathbb{E}(\exp\langle\theta, X\rangle)$, est la transformée de Legendre-Fenchel d'une variable de décision X, $\mathcal{F}(X) : E' \to \overline{\mathbb{R}}$, $\theta \mapsto -\mathbb{V}(-\langle\theta, X\rangle)$ (les signes moins viennent du fait que l'on a considéré le semi-anneau min-plus et non max-plus). Ainsi l'analogie entre optimum et espérance est claire puisque l'optimum d'une variable de décision X est la dérivée en 0 de $\mathcal{F}(X)$ et l'espérance d'une variable aléatoire est la dérivée en 0 de $\mathcal{L}(X)$, et aussi la dérivée en 0 de $\log(\mathcal{L}(X))$.

Lorsque l'optimum d'une variable de décision X à valeurs dans $\mathbb{R}^k$ existe et est unique, et lorsque autour de l'optimum on a le développement limité :

$$\operatorname{conv}(c_X)(x) = \frac{1}{p}\|\sigma^{-1}(x - \mathbb{O}(X))\|^p + o(\|x - \mathbb{O}(X)\|^p) \ ,$$

où $\|\cdot\|$ est la norme quadratique, on dit que X est d'*ordre* p et on appelle σ sa *sensitivité d'ordre* p. Lorsque $p = 2$, σ est l'analogue de l'écart type. Ceci peut se voir en comparant les dérivées secondes en 0 de $\mathcal{F}(X)$ et $\log(\mathcal{L}(X))$.

Une notion reliée est celle de norme p suivante. Si l'optimum d'une variable de décision X à valeurs dans $\mathbb{R}^k$ existe et est unique, on note $|X|_p := \inf\{\sigma \mid c_X(x) \geq \frac{1}{p}(\|(x - \mathbb{O}(X))\|/\sigma)^p\}$ et $\|X\|_p := |X|_p + \|\mathbb{O}(X)\|$. Alors $\|\cdot\|_p$ définit une norme sur l'espace vectoriel $\mathbb{L}^p$ des variables de décision X à valeurs dans $\mathbb{R}^k$ telles que $\|X\|_p < +\infty$, quotienté par l'égalité presque sûre. Ceci permet en particulier de parler de convergence en norme p.

Des variantes des notions introduites ci-dessus ont été introduites et comparées dans les travaux de Bellalouna [Bel92] et de Del Moral [DM94, DM98]. Dans [R1], nous comparons les diverses notions de convergence de variables de décision dans le cas d'un semi-anneau assez général $\mathbb{D}$. Notons en particulier que la convergence en coût implique la convergence presque sûre ce qui est l'implication inverse du cas probabiliste.

Les notions d'optimum, de sensitivité et de norme p permettent d'obtenir des théorèmes limites pour les variables de décision de type loi des grands nombres et théorème de la limite centrale, généralisant ceux établis dans [Qua90].

Théorème 3.2 (Loi des grands nombres, [B5, Theorem 4.6], [R1, Theorem 63])**.** *Si* $(X_n)_{n\in\mathbb{N}}$ *est une suite de variables de décision indépendantes de même mesure de coût (i.i.c.) appartenant à* $\mathbb{L}^p$ *pour* $p \geq 1$*, alors*

$$\lim_{N\to\infty} S_N := \frac{1}{N}\sum_{n=0}^{N-1} X_n = \mathbb{O}(X_0) \ ,$$

où la convergence a lieu en norme p, en coût, presque sûrement et faiblement.

Théorème 3.3 (Théorème de la limite centrale [B5, Theorem 6.6])**.** *Si $(X_n)_{n\in\mathbb{N}}$ est une suite de variables de décision i.i.c. à valeurs dans $\mathbb{R}^k$, d'optimum égal à 0 et de sensitivité d'ordre p égale à σ, et si $1/p+1/p'=1$, alors la suite*

$$Z_N := \frac{1}{N^{1/p'}} \sum_{n=0}^{N-1} X_n$$

converge faiblement vers une variable de décision Z de densité de coût $c_Z(x)=\frac{1}{p}\|\sigma^{-1}x\|^p$.

Il faut noter que, comme indiqué dans les commentaires suivant le théorème 6.6 dans [B5], ce résultat fait appel à un analogue du théorème de Gärtner-Ellis pour les suites de probabilités idempotentes. Dans la section 3.4 nous généralisons ce théorème et sa version idempotente dans un même formalisme.

Comme en théorie des probabilités usuelles, les limites possibles d'un théorème de la limite centrale ne peuvent être que des variables de décision stables, que nous décrivons partiellement dans [C9].

Dans [B5], nous comparons la convergence faible, et la convergence vague de variables de décision avec la convergence en épigraphe de leur densités de coût. Nous montrons aussi un analogue au théorème du porte-manteau pour la convergence faible des probabilités. Nous introduisons, comme en théorie des probabilités usuelles, les notions de tension et de tension asymptotique, et montrons que la tension asymptotique implique la compacité séquentielle pour la convergence faible. Ces résultats peuvent aussi être vus comme des cas particuliers des résultats développés dans le cadre plus général des formes quasi-linéaires présenté dans la section 3.4.

Processus de décision

L'existence de densité des mesures de coût suggère la définition suivante d'une chaîne de Markov par rapport à un espace de décision, que nous appelons ici *chaîne de Bellman*. Si E est un ensemble fini, $C\in\mathbb{R}_{\min}^{E\times E}$ une matrice min-plus stochastique, i.e. telle que $\min_{y\in E} C_{xy}=0$, et ϕ une densité de coût sur E, alors on appelle *chaîne de Bellman* d'espace d'états E, de matrice de transition C et de coût initial ϕ, une variable de décision $X=(X_n)_{n\in\mathbb{N}}$ à valeurs dans $E^{\mathbb{N}}$ telle

$$c_X(x:=(x_0,x_1,\dots))=\phi(x_0)+\sum_{i=0}^{\infty} C_{x_i x_{i+1}} \quad \forall x\in E^{\mathbb{N}} .$$

Une chaîne de Bellman vérifie la propriété de Markov et la densité de coût v^n de la variable de décision X_n vérifie une équation analogue à l'équation de Kolmogorov $v^{n+1}=v^nC$, qui n'est autre ici qu'une équation de Bellman, écrite avec les notations min-plus. Dans [J4], nous avons donné un premier résultat sur les asymptotiques de chaînes de Bellman.

Tout problème de contrôle déterministe non actualisé peut aussi être vu comme l'analogue d'un processus de Markov, que nous appelons ici *processus de Bellman*. Par exemple, un processus de Bellman sur $\mathbb{R}$ à trajectoires continues, est une variable de décision $X=(X_t)_{t\geq 0}$ à valeurs dans l'espace $\mathscr{C}(\mathbb{R}_+)$ des fonctions continues de $\mathbb{R}_+$ dans $\mathbb{R}$ de densité de coût :

$$c_X(x(\cdot)) := \phi(x(0)) + \int_0^{\infty} c(t,x(t),x'(t))dt ,$$

lorsque $x(\cdot)$ est absolument continue et $+\infty$ sinon, où le coût de transition $c : \mathbb{R}_+ \times \mathbb{R}^2 \to \mathbb{R}_{\min}$ vérifie $\inf_{y\in\mathbb{R}} c(t,x,y) = 0$, et ϕ est une densité de coût sur $\mathbb{R}$. On appelle en particulier *processus de décision Brownien d'ordre* p, un processus de Bellman B_t^p à valeurs dans $\mathbb{R}$, tel que $\phi = \mathbb{1}_{\{0\}}$ ($\phi(z) = 0$ si $z = 0$ et $\phi(z) = +\infty$ sinon) et $c(t,x,y) = \frac{1}{p}|y|^p$. Il est donc à incréments indépendants, et tel que $c_{B_t^p - B_s^p}(z) = \frac{|z|^p}{p(t-s)^{p-1}}$.

Dans [B5], nous donnons une condition suffisante de tension dans l'espace $\mathscr{C}([0,1])$ des fonctions continues de $[0,1]$ dans $\mathbb{R}$ [B5, Theorem 8.1]. En utilisant ce résultat de compacité et le théorème de la limite centrale présenté plus haut, on obtient un résultat d'approximation du Brownien min-plus :

Théorème 3.4 ([B5, Theorem 8.3]). *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables de décision i.i.c. à valeurs dans $\mathbb{R}^k$, d'optimum égal à 0, de sensitivité d'ordre p égale à σ, et appartenant à $\mathbb{L}^p$, pour $p \geq 1$, Notons $S_i = X_1 + \cdots + X_i$ les sommes partielles, et $Z^{(n)} = (Z_t^{(n)})_{t\in[0,1]}$ le processus de décision à trajectoires continues défini par :*

$$Z_t^{(n)} = \frac{1}{\sigma n^{1/p'}}(S_{[nt]} + (nt - [nt])X_{[nt+1]}) ,$$

où $1/p + 1/p' = 1$ et où $[\cdot]$ désigne la partie entière. Alors la suite $Z^{(n)}$ converge faiblement vers le processus de décision Brownien d'ordre p.

Une application de ce théorème à l'étude des discrétisations d'équations d'Hamilton-Jacobi est donnée dans la section 4.1.

3.3 Transformée de Cramér

Les travaux présentés ici ont fait l'objet du Rapport de Recherche INRIA [R2].

La *transformée de Cramér* est l'application $\mathcal{C}_{\mathrm{r}}$ qui à une mesure positive μ sur e.v.t.l.c.s. E, de dual E', associe la fonction convexe s.c.i. sur E

$$\mathcal{C}_{\mathrm{r}}(\mu) := \mathcal{F}(\log \mathcal{L}(\mu)) .$$

où $\mathcal{F}$ et $\mathcal{L}$ désignent respectivement les transformées de Legendre-Fenchel sur E' et de Laplace sur E, i.e. $\mathcal{L}(\mu) : E' \to \mathbb{R}_+ \cup \{+\infty\},\ \theta \mapsto \int_E \exp\langle\theta, x\rangle\, d\mu(x)$ et $\mathcal{F}(\phi) : E \to \overline{\mathbb{R}},\ x \mapsto \sup\{< \theta, x > -\phi(\theta) \mid \theta \in E'\}$.

Cette transformation intervient comme la fonction de taux de grandes déviations dans le théorème de Cramér. Si l'on regarde $\mathcal{C}_{\mathrm{r}}(\mu)$ comme une densité de coût sur E, la transformée de Cramér permet d'associer variables aléatoires et variables de décision. Dans [B5], nous décrivons quelques propriétés de cette transformation. Par exemple, la transformée de Cramér envoie un produit de convolution de 2 mesures sur l'inf-convolution de ses transformées, une somme de variables aléatoires indépendantes sur une somme de variables de décision indépendantes, et une loi de probabilité stable sur une densité de coût stable.

Étant donné l'analogie entre principe des grandes déviations, convergence faible de probabilités et convergence faible de variables de décision (ou de mesures de coût), le théorème de Cramér suggère que certains résultats de convergence de variables de décision pourraient être obtenus comme "limites" de convergences de variables aléatoires. De même, les propriétés de la transformée de Cramér suggèrent d'obtenir des convergences de variables de décision en appliquant la transformée de Cramér à des convergences de probabilités. Ceci nécessite alors une propriété de continuité de la transformée de Cramér. Bien sûr, comme l'image de la transformée de Cramér ne contient pas toutes les fonctions de E dans $\overline{\mathbb{R}}$, ni même toutes les fonctions convexes s.c.i., on ne peut espérer obtenir de cette façon des résultats dans la plus grande généralité, mais la connaissance des cas spéciaux pour lesquels cette méthode s'applique peut être utile.

On peut aussi se poser la question inverse, qui est de savoir si les asymptotiques de variables aléatoires pourraient être obtenues comme "limites" d'asymptotiques de variables de décision. Là c'est la non-injectivité de la transformée de Cramér qui élimine l'espoir de résultats généraux, mais les cas spéciaux d'équivalence entre convergence de variables de décision et convergence de variables aléatoires peuvent être utiles dans des travaux ultérieurs.

Dans [R2], nous avons étudié la question de la continuité de la transformée de Cramér. Les résultats d'analyse convexe sur la continuité de la transformée de Legendre-Fenchel pour la convergence en épigraphe, lorsqu'on se restreint à l'ensemble des fonctions convexes s.c.i., et sa non-continuité en général, suggèrent les même propriétés pour la transformée de Laplace. C'est ce que nous prouvons dans [R2, Proposition 21] : le log de la transformée de Laplace est continue de l'espace $\mathcal{M}_{\mathrm{lve}}(E)$ des mesures log-concaves finie non nulles sur E, muni de la topologie de la convergence faible, dans l'espace des fonction convexes s.c.i. propres muni de la topologie de la convergence en épigraphe, pourvu que E vérifie les hypothèses du théorème 3.6 ci-dessous. On donne aussi un contre-exemple pour des probabilités qui ne sont pas log-concaves. Rappelons qu'une mesure positive μ définie sur la tribu de Borel d'un sous-ensemble convexe fermé E d'un e.v.t.l.c.s. X est *log-concave* si

$$\mu(tA + (1-t)B) \geq \mu(A)^t \mu(B)^{1-t} \tag{3.2}$$

pour tous convexes A, B de E et tout $t \in [0,1]$, La continuité de la transformée de Laplace permet ainsi d'obtenir les résultats suivants :

Théorème 3.5 ([R2, Theorem 12])**.** *Si E est un espace vectoriel de dimension finie, la transformée de Cramér est injective et bi-continue de $\mathcal{M}_{\mathrm{lve}}(E)$ dans son image, munis respectivement des topologies de la convergence faible des probabilités usuelles et de la convergence faible des densités de coût.*

Théorème 3.6 ([R2, Theorem 13])**.** *Soit X un e.v.t.l.c.s., et E un convexe fermé de X qui est de plus un espace Polonais (i.e. un espace métrique séparable complet) pour la topologie induite. La transformée de Cramér est alors continue de $\mathcal{M}_{\mathrm{lve}}(E)$ dans son image munis respectivement des topologies de la convergence faible des probabilités usuelles et de la convergence faible des densités de coût.*

3.4 Formes quasi-linéaires et grandes déviations

Les travaux présentés ici ont été obtenus en collaboration avec Stéphane Gaubert (INRIA Rocquencourt), et Vassili Kolokoltsov (Warwick University). Ils ont été présentés dans [C18] et sont prouvés dans la pré-publication [P1].

L'unicité de la solution au problème $(\mathscr{P}')$ étudié dans la section 1.1.2 permet de caractériser la limite en épigraphe de fonctions, la limite de mesures idempotentes, le taux de grandes déviations ou le taux de croissance en temps long de certains problèmes de contrôle stochastique, à partir d'une information plus pauvre donnée par exemple par la limite des transformées de Fenchel-Legendre dans le cas de la convergence en épigraphe, ou la limite des fonctions génératrices dans le cas des grandes déviations. Ceci nous permet en particulier de retrouver et généraliser le théorème de Gärtner-Ellis sur les taux de grandes déviations (voir par exemple [DZ93]).

Formes quasi-linéaires et convergence faible

Dans [C18] nous avons introduit la notion de forme quasi-linéaire et leur convergence faible, ce qui nous permet de regrouper les notions de convergence en épigraphe ou de convergence de probabilités idempotentes, de grandes déviations et de taux de croissance en temps long de problèmes de contrôle stochastique. Nous décrivons brièvement ces notions ici.

On suppose que Y est un espace Polonais. Rappelons que $(\mathbb{R}_{\max})^Y$ est un semi-module idempotent sur $\mathbb{R}_{\max}$. L'addition est alors le supremum pour l'ordre partiel produit $(f,g) \mapsto f \vee g$ et la multiplication d'un élément f de $(\mathbb{R}_{\max})^Y$ par un scalaire $a \in \mathbb{R}_{\max}$ est égale à la fonction $y \in Y \mapsto a + f(y)$ que l'on note $a + f$. On note $\mathrm{sci}(Y)$ (resp. $\mathscr{C}_{\mathrm{b}}(Y)$, resp. $\mathrm{scs}_{\mathrm{b}}(Y)$) l'ensemble des fonctions de Y dans $\mathbb{R}_{\max}$ qui sont s.c.i. (resp. continues bornées supérieurement, resp. s.c.s. bornées supérieurement). Ces ensembles sont tous des sous-semimodules de $(\mathbb{R}_{\max})^Y$.

Définition 3.7. Soit $\mathscr{M}$ un sous-semimodule de $(\mathbb{R}_{\max})^Y$. Une application $F : \mathscr{M} \to \mathbb{R}_{\max}$ (ou $\overline{\mathbb{R}}_{\max}$) est appelée une *forme quasi-linéaire (max-plus)* (sur $\mathscr{M}$) si elle est isotone, additivement homogène, i.e. $F(\lambda + \varphi) = \lambda + F(\varphi)$ pour tout $\lambda \in \mathbb{R}_{\max}$ et $\varphi \in \mathscr{M}$, et si il existe $\alpha \in \mathbb{R}_{\max}$ tel que

$$F(\varphi \vee \psi) \leq \alpha + F(\varphi) \vee F(\psi) \text{ pour tout } \varphi, \psi \in \mathscr{M} \ . \tag{3.3}$$

Une forme quasi-linéaire F sur $\mathscr{M}$ est dite *continue* si elle préserve les suites croissantes convergentes.

On note $\rho(F)$ l'infimum des α vérifiant (3.3), et $\mathscr{QL}(\mathscr{M})$ l'ensemble des formes quasi-linéaires continues de $\mathscr{M}$ dans $\mathbb{R}_{\max}$. Si F prend au moins une valeur finie, alors $\rho(F) \geq 0$ et on peut prendre $\alpha = \rho(F)$ dans (3.3). Sinon $\rho(F) = -\infty$.

Si F est une forme linéaire max-plus alors c'est une forme quasi-linéaire telle que $\rho(F) \leq 0$, et inversement. On note $\mathscr{L}(\mathscr{M})$ l'ensemble des éléments de $\mathscr{QL}(\mathscr{M})$ qui sont linéaires max-plus. Étant donné $f : Y \to \overline{\mathbb{R}}$, l'application $F : (\mathbb{R}_{\max})^Y \to \overline{\mathbb{R}}_{\max}$ définie par

$$F(\varphi) = \sup_{y \in Y} (\varphi(y) - f(y)) \text{ pour tout } \varphi \in (\mathbb{R}_{\max})^Y \tag{3.4}$$

est une forme max-plus linéaire continue (en tant que forme quasi-linéaire) sur $(\mathbb{R}_{\max})^Y$. Une fonction f vérifiant (3.4) est appelée une *densité* de F (l'existence d'une densité peut par exemple être montrée par le théorème 1.1 d'existence du noyau d'une correspondance de Galois fonctionnelle, ou par le théorème 3.1 d'existence de la densité d'une mesure idempotente). Sur $\mathscr{C}_{\mathrm{b}}(Y)$, si F admet une densité, alors il admet une densité minimale qui est s.c.i.

Si μ est une mesure positive finie sur Y, et $\varepsilon > 0$ l'application $F : \mathscr{C}_{\mathrm{b}}(Y) \to \mathbb{R}_{\max}$ définie par

$$F(\varphi) = \varepsilon \log \left(\int_Y \exp \left(\frac{\varphi(y)}{\varepsilon} \right) d\mu(y) \right) \ , \tag{3.5}$$

pour $\varphi \in \mathscr{C}_{\mathrm{b}}(Y)$, est une forme quasi-linéaire continue telle que $\rho(F) \leq \varepsilon \log(2)$. Les applications de cette forme, où μ est une probabilité, apparaissent en théorie des grandes déviations.

Soient $F_i \in \mathscr{QL}(\mathscr{M})$, pour $i \in I$, telles que $\sup_{i \in I} \rho(F_i) < +\infty$. Alors l'application $\sup_{i \in I} F_i$: $\mathscr{M} \to \overline{\mathbb{R}}_{\max}$, $\varphi \mapsto \sup_{i \in I} F_i(\varphi)$ est une forme quasi-linéaire sur $\mathscr{M}$ qui vérifie $\rho(\sup_{i \in I} F_i) \leq$

$\sup_{i\in I}\rho(F_i)$. Ce type de formes apparaît dans les problèmes limites en contrôle stochastique, par exemple les problèmes de contrôle stochastique ergodique.

Dans [C18, P1], on montre que tout élément F de $\mathscr{QL}(\mathscr{C}_{\mathrm{b}}(Y))$ peut être étendu de manière unique sur sci(Y) et admet une extension maximale à $(\mathbb{R}_{\max})^Y$, que l'on note encore F (cette extension ne change pas $\rho(F)$). Ceci permet en particulier de définir une application sur les sous-ensembles de Y encore notée F, par $F(A)=F(\mathbb{1}_A)$ où $\mathbb{1}_A$ est la fonction indicatrice max-plus de A : $\mathbb{1}_A(y)=0$ si $y\in A$ et $\mathbb{1}_A(y)=-\infty$ sinon. Dans le cas où F est linéaire max-plus on obtient une mesure idempotente max-plus (cf. la section 3.1).

On définit la convergence faible et la tension par analogie avec la convergence faible (ou étroite) et la tension de probabilités : si $F\in\mathscr{QL}(\mathscr{C}_{\mathrm{b}}(Y))$ et $(F_n)_{n\in\mathbb{N}}$ est une suite de $\mathscr{QL}(\mathscr{C}_{\mathrm{b}}(Y))$, on dit que F_n *converge faiblement* vers F si $\lim_{n\to\infty}F_n(\varphi)=F(\varphi)$ pour tout $\varphi\in\mathscr{C}_{\mathrm{b}}(Y)$. On dit que $F\in\mathscr{QL}(\mathscr{C}_{\mathrm{b}}(Y))$ est *tendu* si $\inf_{K\subset Y,\,K\text{ compact}}F(K^c)=-\infty$. Notons qu'un élément de $\mathscr{L}(\mathscr{C}_{\mathrm{b}}(Y))$ est tendu si et seulement si sa densité s.c.i. f est inf-compacte, i.e. si $\{y\in Y\mid f(y)\leq\alpha\}$ est compact pour tout $\alpha\in\mathbb{R}$. On définit de manière analogue la *tension asymptotique*. Le résultat suivant est analogue au théorème du porte-manteau pour la convergence faible des probabilités.

Théorème 3.8 ([C18, Theorem 4.1])**.** *Soit $F\in\mathscr{QL}(\mathscr{C}_{\mathrm{b}}(Y))$ et $(F_n)_{n\in\mathbb{N}}$ une suite de $\mathscr{QL}(\mathscr{C}_{\mathrm{b}}(Y))$. Considérons les assertions suivantes :*

$$F_n \text{ converge faiblement vers } F\ , \tag{3.6}$$
$$\liminf_{n\to\infty}F_n(\varphi)\geq F(\varphi) \text{ pour tout } \varphi\in\mathrm{sci}(Y)\ , \tag{3.7}$$
$$\limsup_{n\to\infty}F_n(\varphi)\leq F(\varphi) \text{ pour tout } \varphi\in\mathrm{scs}_{\mathrm{b}}(Y)\ , \tag{3.8}$$
$$\liminf_{n\to\infty}F_n(G)\geq F(G) \text{ pour tout ouvert } G\subset Y\ , \tag{3.9}$$
$$\limsup_{n\to\infty}F_n(C)\leq F(C) \text{ pour tout fermé } C\subset Y\ , \tag{3.10}$$
$$\limsup_{n\to\infty}F_n(K)\leq F(K) \text{ pour tout compact } K\subset Y\ . \tag{3.11}$$

On a (3.7,3.8)$\Rightarrow$(3.6)$\Rightarrow$(3.7)$\Rightarrow$(3.9), (3.8)$\Rightarrow$(3.10)$\Rightarrow$(3.11). *Lorsque $\rho(F)\leq\lim_{n\to\infty}\rho(F_n)=0$, on a de plus* (3.6)$\Leftrightarrow$(3.7,3.8)$\Leftrightarrow$(3.9,3.10). *Si $(F_n)_{n\in\mathbb{N}}$ est asymptotiquement tendue, alors* (3.10)$\Leftrightarrow$(3.11).

Si F_n est de la forme (3.5) pour $\varepsilon=\varepsilon_n$ et $\mu=\mu_n$ où μ_n est une probabilité, F est une forme linéaire max-plus de densité f, et $\lim_{n\to\infty}\varepsilon_n=0$, alors par définition la suite $(\mu_n)_{n\in\mathbb{N}}$ obéit au *principe des grandes déviations* de Varadhan [Var84] pour la fonction de taux f si et seulement si f est inf-compacte ≥ 0 et les conditions (3.9,3.10) sont vérifiées. Dans ce contexte, l'implication (3.9,3.10)$\Rightarrow$(3.6) est appelée le principe de contraction de Varadhan, et d'autres implications sont aussi prouvées dans [Var84] et dans [Puh01, Theorem 3.1.3]. Dans le contexte des capacités les conditions (3.9,3.11) définissent la convergence vague et les conditions (3.9,3.10) définissent la convergence étroite [OV91]. Dans le contexte des formes linéaires max-plus, certaines des implications ont été prouvées dans [B5]. Dans [B5], on montre aussi que les conditions (3.9,3.11) sont équivalentes à la convergence en épigraphe des densités.

On a le résultat de compacité suivant, qui rappelle des résultats analogues dans le contexte des grandes déviations et de la convergence en épigraphe.

Théorème 3.9. *Soit $(F_n)_{n\in\mathbb{N}}$ une suite de $\mathscr{QL}(\mathscr{C}_{\mathrm{b}}(Y))$ telle que $\limsup_{n\to\infty}F_n(Y)<+\infty$ et $\lim_{n\to\infty}\rho(F_n)=0$. Il existe $F\in\mathscr{L}(\mathscr{C}_{\mathrm{b}}(Y))$ et une sous-suite de $(F_n)_{n\in\mathbb{N}}$ vérifiant les conditions* (3.9,3.11).

Un théorème de Gärtner-Ellis pour les formes quasi-linéaires

Soient $X, Y, B, B^\circ, b, b^\circ$ comme dans la section 1.1.2 et tels que B est une conjugaison de Moreau. On notera $\bar{b}$ le noyau de la conjugaison de Moreau, tel que $b(x, y, \lambda) = \bar{b}(x, y) - \lambda$.

De même on note $\bar{b}_{x,V}(y) = \sup_{z\in V} \bar{b}(z, y) - \bar{b}(x, y)$. La fonction $b^\alpha_{x,V}$ de (1.3) vérifie donc $b^\alpha_{x,V}(y) = \bar{b}_{x,V} + \alpha$.

On dit que b (ou $\bar{b}$) est *fortement coercive* si pour tout $x \in X$ et tout voisinage V de x dans X, il existe un sous-ensemble fini W de V tel que la fonction $\bar{b}_{x,W}$ a ses ensembles de sous-niveau finis relativement compacts.

On dit que $\bar{b}$ est *supérieurement (fortement) coercive* si pour tout $x \in X$, et tout voisinage V de x dans X, il existe un sous-ensemble fini W de V tel que $\bar{b}(x, \cdot)$ est bornée supérieurement sur tout ensemble de sous-niveau fini de $\bar{b}_{x,W}$.

Le noyau de la transformée de Legendre-Fenchel est toujours supérieurement coercif, et il est fortement coercif sur $\mathbb{R}^n$.

Le résultat suivant justifie l'étude du problème $(\mathscr{P}')$.

Théorème 3.10 ([C18, Theorem 5.1]). *Soit $(F_n)_{n\in\mathbb{N}}$ une suite de $\mathscr{QL}(\mathscr{C}_\mathrm{b}(Y))$ telle que $\lim_{n\to\infty} \rho(F_n) = 0$. Supposons que F_n converge faiblement vers $F \in \mathscr{L}(\mathscr{C}_\mathrm{b}(Y))$, de densité s.c.i. f, et que b est continue en la seconde variable et supérieurement coercive. Considérons la fonction $g : X \to \overline{\mathbb{R}}$ définie par :*

$$g(x) = \limsup_{n\to\infty} F_n(b(x, \cdot)) \quad \textit{pour tout } x \in X \ . \tag{3.12}$$

Alors

$$Bf \leq g \textit{ et } Bf = g \textit{ sur } \mathrm{idom}(g) \cup g^{-1}(-\infty) \ . \tag{3.13}$$

Ainsi, si le système (3.13) a une unique solution f, et si la lim sup dans (3.12) est une limite, toutes les limites faibles de sous-suites de la suite $(F_n)_{n\in\mathbb{N}}$ possibles sont égales à la forme max-plus linéaire F dont la densité est l'unique solution f de $(\mathscr{P}')$. Pour obtenir la convergence de la suite, il suffit alors de vérifier sa compacité relative, ce qui est vrai si la suite est asymptotiquement tendue. Toutes ces propriétés sont assurées sous les hypothèses du théorème suivant qui généralise le théorème de Gärtner-Ellis.

Théorème 3.11 ([C18, Corollary 5.4]). *Soit $(F_n)_{n\in\mathbb{N}}$ une suite de $\mathscr{QL}(\mathscr{C}_\mathrm{b}(Y))$ telle que $\lim_{n\to\infty} \rho(F_n) = 0$, Définissons $g : X \to \overline{\mathbb{R}}$ par* (3.12), *et supposons que la limsup est une limite. Supposons* (H6) *ou* (H6)$'$ *vérifiée, b fortement coercive, et $B^\circ g$ quasi-continue dans son domaine. Supposons qu'il existe $x_0 \in \mathrm{idom}(g)$ tel que $\bar{b}(x_0, \cdot)$ est bornée inférieurement par une constante. Alors $\{(\partial^\circ g)^{-1}(y)\}_{y\in\mathrm{ldom}(B^\circ g)}$ est un recouvrement de* $\mathrm{idom}(g)$. *De plus, si ce recouvrement est topologiquement minimal, alors F_n converge faiblement vers la forme linéaire max-plus de densité $B^\circ g$.*

D'autres résultats sont montrés dans [C18, P1] qui généralisent en particulier des formes plus précises du théorème de Gärtner-Ellis telles que celles présentées dans [DZ93]. Ces résultats permettent aussi d'obtenir des théorèmes limites en contrôle stochastique.

3.5 Perspectives / Travaux en cours

Dans [BCJ00], Barron et Cardaliaguet ont établi un résultat d'existence de densité d'une mesure idempotente par rapport à une mesure usuelle. Ce résultat n'est pas comparable avec ceux des sections 1.1.1 et 3.1. Dans son travail de mémoire de M2, co-encadré par F. Baccelli et moi-même, Paul Poncet a obtenu un théorème de Radon-Nikodým généralisant à la fois le résultat d'existence de densité de la section 3.1 et celui de [BCJ00], au moyen d'un résultat de séparation de convexes max-plus établi dans [CGQ04]. Ce mémoire contient aussi l'étude et la comparaison des diverses notions de mesures idempotentes apparues dans plusieurs domaines des mathématiques.

Plusieurs directions de recherche relatives aux probabilités idempotentes et aux grandes déviations restent encore à développer.

Comme indiqué dans l'introduction, la bijection $x \mapsto -\epsilon \log x$ envoie le semi-corps $(\mathbb{R}_+, +, \times)$ dans semi-corps $\mathbb{R}_\epsilon$ dont la limite quand $\epsilon \to 0^+$ n'est autre que l'algèbre min-plus. On peut ainsi transporter l'analyse classique et la théorie des probabilités de $\mathbb{R}_+$ vers $\mathbb{R}_\epsilon$, et obtenir les mesures de coût, par passage à la limite quand $\alpha \to 0$. En particulier, en regroupant dans un même ensemble les probabilités sur les semi-corps $\mathbb{R}_\epsilon$ et les mesures de coût, on devrait pouvoir développer une approche globale probabilités–optimisation, intégrant en particulier mesures de coût et probabilités usuelles, tout en conservant les propriétés induites par la structure algébrique. Sur cet ensemble, la notion d'intégrale peut être définie de manière non ambiguë, ce qui n'est pas le cas en général dans l'espace plus grand des capacités ou même dans celui des formes quasi-linéaires développées dans la section 3.4 restreintes aux fonctions indicatrices d'ensembles. Ce point de vue permettrait donc d'algébriser une partie de la théorie des grandes déviations à la loi des grands nombres, et d'obtenir des résultats plus précis que ceux de la section 3.4.

Les travaux sur les densités de coût stables, et les asymptotiques de chaînes de Bellman, n'ont pas été développés. On pourrait chercher en particulier des outils de calcul simples à partir de mesures de coût typiques dans l'esprit de la théorie des probabilités, en vue de calculer ou d'étudier les limites de mesures de coût mais aussi les taux des grandes déviations. Les travaux de Puhalskii [Puh01] et de Fleming [Fle04] sur les maxingales ou martingales max-plus peuvent être utiles dans cette perspective.

L'étude d'applications concrètes en contrôle stochastique (optimisation de portefeuille,...) ou en probabilités (systèmes de particules, de files d'attente,...) devrait pouvoir bénéficier des résultats de la section 3.4.

Finalement, au cours des travaux sur les asymptotiques de valeurs propres et vecteurs propres présentés dans le chapitre 5, nous avons été amenés à utiliser un semi-anneau de jets d'ordre 1 (qui sont les développements asymptotiques du 1er ordre), afin d'obtenir facilement des asymptotiques de type grandes déviations précises. On pourrait de même essayer de développer une théorie des probabilités sur ce semi-anneau de jets (qui n'est ni symétrisable, ni idempotent), comme outil de calcul de grandes déviations précises.

Chapitre 4

Méthodes max-plus en analyse numérique de problèmes de contrôle déterministe

Considérons le problème de contrôle déterministe à temps et espace continus et à horizon fini $t > 0$ (4a) défini dans l'introduction, où le supremum est pris sur toutes les trajectoires $(\mathbf{x}(\cdot), \mathbf{u}(\cdot))$ telles que $\mathbf{u}(\cdot)$ est mesurable à valeurs dans l'ensemble $U \subset \mathbb{R}^m$ des contrôles, $\mathbf{x}(\cdot)$ est absolument continu à valeurs dans l'espace $\Omega \subset \mathbb{R}^n$ des états, et vérifiant (4b) pour une condition initiale $x \in \Omega$ donnée. Si le *gain instantané* ou *lagrangien* $\ell : \Omega \times U \to \mathbb{R}$, la *dynamique* $g : \Omega \times U \to \mathbb{R}^n$, et le *gain final* $\phi : \Omega \to \mathbb{R} \cup \{-\infty\}$ sont assez réguliers, la fonction valeur $v : [0, +\infty[\times\Omega \to \overline{\mathbb{R}}$ est l'unique solution de viscosité de l'équation de la programmation dynamique d'Hamilton-Jacobi (5). D'autre part, comme rappelé dans l'introduction, le semi-groupe d'évolution S^t de l'équation d'Hamilton-Jacobi ou du problème de contrôle, qui associe à ϕ la fonction $v^t = v(t, \cdot) : \Omega \to \overline{\mathbb{R}}$ est max-plus linéaire : $S^t(f \vee g) = S^t(f) \vee S^t(g)$ et $S^t(\lambda + f) = \lambda + S^t(f)$.

On cherche ici à développer des méthodes numériques pour résoudre l'équation d'Hamilton-Jacobi (5), ou de manière équivalente le problème de contrôle (4), qui tiennent compte des propriétés du problème : déterminisme et linéarité max-plus. Rappelons que les discrétisations classiques qui préservent la monotonie, par exemple les différences finies décentrées (voir par exemple [KD92]), ou les discrétisations semi-lagrangiennes (voir par exemple [CDF89]) conduisent toutes à rajouter de la viscosité artificielle : l'équation approchée est l'équation de la programmation dynamique d'un problème de contrôle optimal stochastique. On voudrait développer des discrétisations en temps et en espace de l'équation d'Hamilton-Jacobi (5), ou du problème de contrôle (4), qui conduisent à la résolution de l'équation de la programmation dynamique d'un problème de contrôle optimal déterministe en temps et espace discret. Cette équation serait alors max-plus linéaire. Les travaux de Fleming et McEneaney (voir par exemple [FM00]) vont dans cette direction.

Dans la section 4.1, nous montrons brièvement comment le théorème d'approximation du processus de décision Brownien établi dans la section 3.2 implique qu'une discrétisation conduisant à une équation max-plus linéaire ne peut être locale.

Dans la section 4.2, nous présentons les travaux de thèse d'Asma Lakhoua, dans lesquels une méthode des éléments finis max-plus a été introduite et étudiée. Cette méthode conduit à une équation de la programmation dynamique d'un jeux à somme nulle déterministe, alors que celle de Fleming et McEneaney conduit à une équation max-plus linéaire, qui s'interprète donc comme l'équation de la programmation dynamique d'un problème de contrôle optimal (i.e. à un joueur) déterministe. Elle a l'avantage de permettre des estimations d'erreurs systématiques en termes de projecteurs sur des semi-modules max-plus de fonctions.

4.1 Approximation des processus de Bellman

Nous traduisons ici le théorème 3.4 d'approximation du processus de décision Brownien d'ordre p en termes de fonctions valeurs de problèmes de contrôle. Le théorème nous dit que si f est une fonction continue bornée supérieurement sur $\mathscr{C}([0,1])$, alors $\lim_{n\to\infty} \mathbb{V}(f(Z^{(n)})) = \mathbb{V}(f(B^p))$. Rappelons que si X est une variable de décision de densité de coût c_X, la valeur de $f(X)$ s'écrit $\mathbb{V}(f(X)) = \inf_x f(x) + c_X(x)$. Considérons les fonction $f_{t,\xi}$ définies par

$$f_{t,\xi}(x(\cdot)) = \int_0^t \psi(x(s)+\xi)\,ds + \phi(x(t)+\xi) \quad .$$

Alors

$$v(t,\xi) := \mathbb{V}(f_{t,\xi}(B^p)) = \inf_{x(\cdot),x(0)=\xi} \int_0^t \left(\psi(x(s)) + \frac{1}{p}|x'(s)|^p\right) ds + \phi(x(t))$$

est la fonction valeur d'un certain problème de contrôle. D'autre part si c est la densité de X_1 et si $\sigma = 1$ et $t = k/n$, alors

$$v_n(t,\xi) := \mathbb{V}(f_{t,\xi}(Z^{(n)})) = \inf_{x(\cdot),x(0)=\xi} \left[\int_0^t \psi(x(s))\,ds + \sum_{i=1}^k c(n^{1/p'}(x(i/n) - x((i-1)/n))) + \phi(x(t))\right]$$

où on restreint l'infimum aux trajectoires continues x qui sont affines sur les intervalles $[k/n,(k+1)/n]$. Ainsi, v_n est une approximation en temps de v, qui vérifie l'équation de la programmation dynamique :

$$v_n(t+1/n,x) = \inf_y \left(\int_0^{1/n} \psi(x + s\,n(y-x))\,ds + c(n^{1/p'}(y-x)) + v_n(t,y)\right), \quad t \in \frac{1}{n}\mathbb{N} \ .$$

Elle serait donc aussi une approximation en espace si la fonction c était à support fini, i.e. était égale à $+\infty$ sauf en un nombre fini de points. Mais les conditions du théorème demandent en particulier que $\mathrm{conv}(c)(x) = \frac{1}{p}|x|^p + o(|x|^p)$ autour de 0, ce qui est impossible pour une fonction à support fini. En fait, le résultat de convergence demeure pour une suite de variables de décision $(X_m)_{m\geq 0}$ dépendant aussi du pas de discrétisation en temps $1/n$, pourvu que la densité c_n de X_1 vérifie $\lim_{n\to\infty} n\mathrm{conv}(c_n)(\frac{x}{n^{1/p}}) = \frac{1}{p}|x|^p$. On peut alors choisir c_n à support fini pour tout n, mais on voit bien que pour obtenir la courbure de $\frac{1}{p}|x|^p$, il faut que le support de c_n contienne de plus en plus de points autour de 0.

Ceci montre en particulier qu'une approximation d'un problème de contrôle déterministe par un problème de contrôle déterministe à temps discret et espace d'état fini, ne peut pas être à transitions locales, i.e. tel que pour tout x, le coût de transition $c(x,y)$ soit nul sauf sur un ensemble fini borné de points autour de x.

4.2 Méthode des éléments finis max-plus

Les travaux présentés ici font partie du travail de thèse d'Asma Lakhoua. Cette thèse est en co-tutelle Paris VI–ENIT, elle est coencadrée par Stéphane Gaubert (directeur de thèse pour Paris VI), Henda El Fekih (LAMSIN, directeur de thèse ENIT, Tunis) et moi-même. Le travail de thèse a fait l'objet des articles publiés ou acceptés pour publication [C16, C17, J16].

On s'intéresse ici à l'équation (5), pour laquelle on développe un analogue max-plus de la méthode des éléments finis de Petrov-Galerkin. Pour cela, on remplace (5a) par la "formulation variationnelle" suivante suggérée par la linéarité du semi-groupe S^t, et introduite par Kolokoltsov et Maslov (voir par exemple [KM97, Section 3.2]) :

$$v^{t+\delta} \in \mathcal{W}, \quad \langle z, v^{t+\delta} \rangle = \langle z, S^\delta v^t \rangle \quad \forall z \in \mathcal{Z} \ , \tag{4.1}$$

pour tous $t \geq 0, \delta > 0$, où $\mathcal{W}$ est un semi-module complet de fonctions de Ω dans $\overline{\mathbb{R}}_{\max}$ contenant tous les fonctions v^t, et où $\mathcal{Z}$ est un semi-module "dual" de "fonctions tests" de Ω dans $\overline{\mathbb{R}}_{\max}$. Ici $\langle \cdot, \cdot \rangle$ désigne le produit scalaire max-plus, i.e. $\langle u, v \rangle = \sup_{x \in \Omega} u(x) + v(x)$.

La discrétisation en temps de pas $\delta > 0$ de (4.1) revient à ne vérifier cette équation que pour ce δ et pour $t = 0, \delta, \ldots, T - \delta$, où T est un horizon.

La discrétisation en espace consiste à remplacer respectivement les semi-modules $\mathcal{W}$ et $\mathcal{Z}$ par des sous-semi-modules $\mathcal{W}_h$ et $\mathcal{Z}_h$ finiment engendrés. Comme l'équation qui en résulte n'a pas forcément une solution, on définit $v_h^{t+\delta}$ comme étant la sous-solution maximale de ce système d'équations, i.e. la solution maximale $v_h^{t+\delta} \in \mathcal{W}$ aux inégalités : $\langle z, v_h^{t+\delta} \rangle \leq \langle z, S^\delta v_h^t \rangle$ pour tous $z \in \mathcal{Z}_h$. Si $\mathcal{W}_h$ est généré par la famille $\{w_i\}_{1 \leq i \leq p}$ d'éléments finis, et $\mathcal{Z}_h$ est généré par la famille $\{z_j\}_{1 \leq j \leq q}$ de fonctions test, et si on note W_h l'opérateur max-plus linéaire de $\overline{\mathbb{R}}_{\max}^p$ dans $\mathcal{W}$ tel que $W_h \lambda = \sup_{1 \leq i \leq p} w_i \lambda_i$ pour tout $\lambda = (\lambda_i)_{i=1,\ldots,p} \in \overline{\mathbb{R}}_{\max}^p$, on obtient [J16, Proposition 4] que la solution maximale $\lambda^t \in \mathbb{R}_{\max}^p$ à l'équation $v_h^t = W_h \lambda^t$ pour tout $t = 0, \delta, \cdots, T$, suit le système dynamique

$$\lambda^{t+\delta} = M_h^\sharp (K_h \lambda^t) \ , \tag{4.2}$$

pour $t = 0, \cdots, T - \delta$, avec la condition initiale : $\lambda^0 = W_h^\sharp \phi$, et où M_h et K_h sont des matrices max-plus $q \times p$ définies par :

$$(M_h)_{ji} = \langle z_j, w_i \rangle, \quad (K_h)_{ji} = \langle z_j, S^\delta w_i \rangle \ . \tag{4.3}$$

En notations usuelles, l'équation (4.2) s'écrit :

$$\lambda_i^{t+\delta} = \min_{1 \leq j \leq q} \left(-(M_h)_{ji} + \max_{1 \leq k \leq p} \left((K_h)_{jk} + \lambda_k^t \right) \right) \quad \text{pour } 1 \leq i \leq p \ .$$

Elle s'interprète donc comme l'équation de la programmation dynamique d'un problème de jeux à somme nulle déterministe.

La méthode proposée initialement par Fleming et McEneaney dans [FM00] consiste à approcher l'équation $v^{t+\delta} = S^\delta v^t$ par une équation linéaire max-plus sur le semi-module $\mathcal{W}_h$ généré par un nombre fini de fonctions de "base". Cela revient à considérer $\mathcal{Z}_h = \mathcal{Z}$, et à approcher v^t par $v_h^t = W_h \mu^t$, avec $\mu^{t+\delta} = S_h^\delta \mu^t$, où S_h est la matrice max-plus maximale telle que $M_h S_h \leq K_h$. On en déduit que $W_h \mu^t \leq W_h \lambda^t \leq v^t$ [J16, Proposition 7].

Une autre interprétation de la récurrence (4.2) peut être obtenue en termes de projections puisque $v_h^{t+\delta} = S_h^\delta v_h^t$, pour $S_h^\delta := P_{\mathcal{W}_h}^{\mathcal{Z}_h} \circ S^\delta$, où $P_{\mathcal{W}_h}^{\mathcal{Z}_h}$ est un projecteur, égal à la composée $P_{\mathcal{W}_h} \circ P^{-\mathcal{Z}_h}$ d'un projecteur (par le dessous) sur le semi-module max-plus $\mathcal{W}_h$ ($P_{\mathcal{W}_h}(u) = \sup\{w \in \mathcal{W}_h \mid w \leq u\}$) et d'un projecteur (par le dessus) sur le semi-module min-plus $-\mathcal{Z}_h$ ($P^{-\mathcal{Z}_h}(u) = -P_{\mathcal{Z}_h}(-u)$). Dans [J16], on obtient une estimation générale de l'erreur au moyen des erreurs dues aux projections des fonctions v^t, $t = 0, \delta, \dots, T$, sur $\mathcal{W}_h$ et $-\mathcal{Z}_h$. Selon la régularité de la fonction valeur (semi-convexité, continuité lipschitzienne), et la régularité des éléments finis (Diracs, normes, formes quadratiques), on peut obtenir une erreur de projection de l'ordre de Δx ou $(\Delta x)^2$ qui induit une contribution à l'erreur totale de l'ordre de $\frac{\Delta x}{\delta}$ ou $\frac{(\Delta x)^2}{\delta}$.

La méthode présentée ici nécessite le calcul de $\langle z, S^\delta w\rangle$ pour toute fonction test z et tout élément fini w. Plusieurs approximations ont été proposées et étudiées tant théoriquement que numériquement dans [C16, C17, J16]. On peut par exemple approcher $S^\delta w$ par $w + \delta H(x, \frac{\partial w}{\partial x})$, ce qui peut être raisonnable lorsque w est régulière. Les approximations proposées ajoutent un terme de l'ordre de δ ou $\sqrt{\delta}$ à l'erreur totale, selon le type d'approximation choisi.

La récurrence (4.2) fait appel à des produits de matrices max-plus ou min-plus pleines, ce qui rend l'implémentation difficile. Les résultats récents de la thèse d'Asma Lakhoua montrent que l'on peut se ramener à des matrices semi-creuses. Il faut toutefois noter que, de même que dans le cas d'approximations max-plus linéaires décrites dans la section précédente, il est impossible de se contenter d'une approximation par des matrices creuses.

La méthode des éléments finis peut s'appliquer aussi à des problèmes de contrôle déterministe à horizon infini actualisé. Dans ce cas, on est ramené à la recherche du point fixe de l'opérateur de la programmation dynamique d'un problème de jeux à somme nulle, que l'on peut résoudre au moyen d'un algorithme de type itération sur les politiques. L'actualisation et la contraction au sens large des opérateurs max-plus linéaires implique la contraction stricte de l'opérateur de la programmation dynamique et les techniques utilisées dans le cas à horizon fini s'adaptent.

4.3 Perspectives / Travaux en cours

Les estimations de convergence prouvées dans [J16] sont obtenues pour la norme infinie, et ne s'appliquent que lorsque la fonction valeur est lipschitzienne. On pourrait envisager d'obtenir des résultats de convergence plus faible sous des hypothèses plus faibles de régularité. On pourrait par exemple utiliser les notions de convergence (en coût, faible, vague,...) introduites dans [B3, R1, B5] (voir la section 3.2), ou les techniques de solutions de viscosité à la Crandall-Ishii-Lions.

La résolution de problèmes de contrôle ergodique (i.e. sans taux d'actualisation), conduit par discrétisation à la résolution d'un problème de jeux à somme nulle ergodique, que l'on pourrait résoudre au moyen de l'algorithme de type itérations sur les politiques introduit dans [CTG06]. Néanmoins l'analyse du schéma est rendue délicate par le fait qu'on a seulement la contraction au sens large des opérateurs.

Chapitre 5

Perturbation et calcul de valeurs propres

Les travaux présentés ici ont été obtenus en collaboration avec Ravindra Bapat (Indian Statistical Institute, New Delhi) et Stéphane Gaubert (INRIA Rocquencourt). Ils ont fait l'objet des articles publiés [J5, C15, J14], et de la pré-publication [S1].

Soit $\mathcal{A}_\epsilon$ une matrice $n \times n$ dépendant continuement du paramètre $\epsilon > 0$ et telle que les limites

$$A_{ij} := \lim_{\epsilon \to 0^+} \frac{\log(\mathcal{A}_\epsilon)_{ij}}{\log \epsilon} \tag{5.1}$$

existent pour tous $i, j = 1, \ldots, n$. Les remarques de l'introduction suggèrent que les valeurs propres et vecteurs propres de $\mathcal{A}_\epsilon$ ont des asymptotiques du même type, et que leurs limites sont les valeurs propres et vecteurs propres min-plus de la matrice des limites $A = (A_{ij})$. De même si les coefficients de $\mathcal{A}_\epsilon$ admettent un développement en série de Puiseux par rapport à ϵ, on sait que les valeurs propres et vecteurs propres admettent eux aussi un développement en série de Puiseux par rapport à ϵ, et on peut espérer calculer leurs exposants au moyen de valeurs propres et de vecteurs propres de matrices min-plus construites à partir des exposants des développements de Puiseux des coefficients de $\mathcal{A}_\epsilon$.

Nous cherchons ici à calculer les asymptotiques ou les exposants de développements de Puiseux des valeurs propres et vecteurs propres au moyen d'outils d'algèbre min-plus. Concernant le développement de Puiseux des valeurs propres de $\mathcal{A}_\epsilon$, une manière naturelle serait de calculer le polynôme caractéristique de cette matrice, et de calculer ensuite les exposants des racines de ce polynôme au moyen de l'algorithme de Puiseux. Ce calcul lourd perd malheureusement la structure matricielle sous-jacente, que nous cherchons au contraire à utiliser. Il est par contre utile dans les preuves des résultats recherchés.

Dans la section 5.1, nous étudions la valeur et le vecteur propre de Perron d'une matrice à coefficients positifs ou nuls. Les sections suivantes concernent les développement au premier ordre des valeurs propres de matrice à coefficients complexes. Comme les matrices min-plus irréductibles n'ont qu'une valeur propre possible, la difficulté est de définir des “valeurs propres”, qui seraient des candidats aux exposants de toutes les valeurs propres de $\mathcal{A}_\epsilon$. Dans la section 5.3, nous généralisons un théorème de Višik, Ljusternik, Lidskiĭ sur les perturbations de valeurs propres, au moyen de valeurs propres min-plus de compléments de Schur de matrices. Dans la section 5.4, nous considérons plus généralement des faisceaux de matrices, et utilisons les “racines” du polynôme caractéristique d'un faisceau de matrices min-plus, défini à l'aide du permanent min-plus, qui n'est autre qu'un problème d'affectation optimale.

Les résultats présentés ici montrent donc que les problèmes de perturbations de valeurs propres sont gouvernés par des problèmes d'optimisation combinatoire.

5.1 Asymptotiques de la valeur propre et du vecteur propre de Perron

Considérons une matrice $\mathcal{A}_\epsilon$ à coefficients réels positifs ou nuls, irréductible. On sait alors qu'elle admet une unique valeur propre $\mathcal{L}_\epsilon \in \mathbb{R}_+$ associée à un vecteur propre $\mathcal{U}_\epsilon \in (\mathbb{R}_+)^n \setminus 0$. La valeur propre, qui est aussi égale au rayon spectral $\rho(\mathcal{A}_\epsilon)$ de $\mathcal{A}_\epsilon$, est appelée *valeur propre de Perron* de A, et le vecteur propre $\mathcal{U}_\epsilon$ est unique à constante multiplicative près, et est appelé *vecteur propre de Perron* de A. On supposera par la suite que ce vecteur est normalisé, i.e. qu'il vérifie $\sum_{i=1}^n (\mathcal{U}_\epsilon)_i = 1$.

Si l'on pose $T = -1/(\log \epsilon)$, alors T tend vers 0^+, lorsque ϵ tend vers 0^+, et on peut voir (5.1) comme une asymptotique de type grandes déviations. Le but ici est de trouver ce type d'asymptotiques pour la valeur propre et le vecteur propre de Perron. Ce problème apparaît en particulier en mécanique statistique dans la méthode de l'opérateur de transfert à petite température T. Le cas où $\mathcal{A}_\epsilon$ est une matrice de Markov, et où l'on cherche les asymptotiques du vecteur propre à gauche, est déjà traité par la théorie de Freidlin-Wentzell.

En utilisant le théorème spectral min-plus (dont le théorème 2.1 est une généralisation au cas de matrices dénombrables max-plus), on obtient facilement [J5, Théorème 1] que si les limites (5.1) existent dans $\mathbb{R}_{\min}$ et si $A = (A_{ij})$ est irréductible, alors $\lim_{\epsilon\to 0^+} \frac{\log \mathcal{L}_\epsilon}{\log \epsilon}$ existe et est égale à la valeur propre min-plus $\rho_{\min}(A)$ de A (i.e. le minimum des poids moyens des circuits de A). Si de plus A n'a qu'une classe critique, alors $\lim_{\epsilon\to 0^+} \frac{\log(\mathcal{U}_\epsilon)_i}{\log \epsilon} = U_i$ où $U = (U_i)$ est l'unique vecteur propre de A dans $\mathbb{R}_{\min}$ vérifiant la normalisation $\sum_{i=1}^n U_i = \mathbb{1}$ [J5, Théorème 2]. Ce vecteur est proportionnel à n'importe quelle colonne de $\tilde{A}^\star$ d'indice critique.

Afin d'obtenir des asymptotiques de type grandes déviations précises, on considère l'algèbre des fonctions f de $\epsilon > 0$ admettant un développement limité autour de $\epsilon = 0^+$ de la forme $f(\epsilon) = b\epsilon^B + o(\epsilon^B)$ avec $b > 0$ et $B \in \mathbb{R}$, ou avec $b = 0$ et $B = +\infty$ (où par convention $0\epsilon^{+\infty} \equiv 0$). Cette algèbre est munie de l'addition et de la multiplication usuelle, et on appelle semi-corps de *jets*, et on note $\mathbb{J}_{\min}$, l'image de cette algèbre par l'application $\pi : f \mapsto (b, B)$. Le semi-corps $\mathbb{J}_{\min}$ a pour zéro $(0, +\infty)$ et pour unité $(1, 0)$, notés simplement 0 et 1. On écrit simplement $f(\epsilon) \sim (b, B)$ lorsque $\pi(f) = (b, B)$. On étend la notation $\sim$ aux vecteurs et matrices (coordonnée par coordonnée).

Rappelons qu'une classe basique d'une matrice à coefficients positifs, est telle que le rayon spectral de cette matrice restreinte à cette classe est maximal. Si $a \in \mathbb{R}_+^{n\times n}$ et G est un graphe, on note a^{G} la matrice obtenue à partir de a en annulant les coefficients a_{ij} pour lesquels l'arc (i,j) n'est pas dans G. Si $A \in \mathbb{R}_{\min}^{n\times n}$ et $U \in \mathbb{R}_{\min}^n$, on appelle *graphe de saturation* de U par rapport à A, le graphe formé des nœuds $1, \ldots, n$ et d'un arc (i,j) lorsque $A_{ij}U_j = (AU)_i$. Le résultat suivant est à la fois un théorème spectral dans $\mathbb{J}_{\min}$, et un résultat d'asymptotiques (ou grandes déviations précises) pour les valeurs propres et vecteurs propres de Perron :

Théorème 5.1 ([J5, Théorème 3])**.** *Si $\mathcal{A}_\epsilon \sim \mathcal{A} = (a, A) \in (\mathbb{J}_{\min})^{n\times n}$, avec $\mathcal{A}$ irréductible, alors $\mathcal{L}_\epsilon \sim \rho_{\mathbb{J}}(\mathcal{A})$ où $\rho_{\mathbb{J}}(\mathcal{A}) = (\rho(a^{\mathsf{GC}}), \rho_{\min}(A))$ est la valeur propre de $\mathcal{A}$ dans $\mathbb{J}_{\min}$, et GC est le graphe critique de A.*

Si a^{GC} n'a qu'une classe basique, alors $\mathcal{U}_\epsilon \sim \mathcal{U}$, où $\mathcal{U}$ est l'unique vecteur propre de $\mathcal{A}$ dans $\mathbb{J}_{\min}^n$

de somme 1. *Celui-ci est de la forme* (u, U)*, où* U *est une colonne de* $\tilde{A}^\star$ *d'indice basique et* u *est un vecteur propre positif de la matrice* a^{S}*, où* S *est le graphe de saturation de* U *par rapport à* A*.*

Si a^{GC} admet $s \geq 2$ classes basiques, alors $\mathcal{A}$ admet plusieurs vecteurs propres dans $\mathbb{J}^n_{\min}$, ceux-ci étant déterminés linéairement par la donnée d'un vecteur de $\mathbb{J}^s_{\min}$. Les asymptotiques de $\mathcal{U}$ ne peuvent alors être déterminées qu'au moyen des termes suivants du développement de Puiseux de $\mathcal{A}_\epsilon$. Dans [J5], on a proposé une procédure d'agrégation à la Freidlin-Wentzell. Celle-ci calcule par ailleurs le terme suivant du développement limité de la valeur propre de Perron.

Pour aller plus loin, il faut abandonner le cadre des matrices à coefficients positifs, et passer au cas complexe, ce que nous faisons dans la section suivante.

5.2 Exposants de racines et racines min-plus

Considérons maintenant une matrice $\mathcal{A}_\epsilon$ admettant un développement limité autour de 0^+ de la forme

$$(\mathcal{A}_\epsilon)_{ij} = a_{ij}\epsilon^{A_{ij}} + o(\epsilon^{A_{ij}}) \ , \tag{5.2}$$

avec $a_{ij} \in \mathbb{C}$ et $A_{ij} \in \mathbb{R}_{\min}$ (où l'on pose $a\epsilon^A \equiv 0$ si $a = 0$ ou $A = +\infty$). Si $A \in \mathbb{R}^{n\times n}_{\min}$ est une matrice min-plus, on note $\operatorname{per} A$ son permanent, qui est la somme min-plus (c'est-à-dire le minimum) sur toutes les permutations σ des poids $A_{1\sigma(1)} \cdots A_{n\sigma(n)}$. C'est donc aussi la valeur d'une affectation optimale dans le graphe associé à A, si A_{ij} représente le coût de l'affectation de j à i. On appelle polynôme caractéristique de A le polynôme formel $\operatorname{per}(\mathsf{Y}I \oplus A)$. Étant donné un polynôme formel min-plus $P(\mathsf{Y}) = \bigoplus_{j=0}^n P_j \mathsf{Y}^j$, on sait, d'après un résultat de Cuninghame-Green et Meijer (voir [BCOQ92]) que sa fonction polynomiale associée $\widehat{P} : \mathbb{R}_{\min} \to \mathbb{R}_{\min}$ se factorise de manière unique sous la forme $\widehat{P}(y) = P_n(y \oplus c_1) \cdots (y \oplus c_n)$, avec $c_1 \leq \cdots \leq c_n \in \mathbb{R}_{\min}$. En général cette factorisation n'a pas lieu pour le polynôme formel lui-même, sauf s'il est convexe (dans le sens où la fonction $k \mapsto P_k$ est égale à la restriction à $\{0, \ldots, n\}$ d'une fonction convexe sur $\mathbb{R}$). Nous appelons $c_1, \ldots, c_n$ les *racines* de P (comptées avec leur multiplicités).

En appliquant une version légèrement généralisée dans [S1, Theorem 3.1] du théorème de Newton-Puiseux pour le calcul du premier terme du développement de Puiseux des racines du polynôme caractéristique de $\mathcal{A}_\epsilon$, on obtient :

Théorème 5.2 ([S1, Theorem 3.8])**.** *Soit* $\mathcal{A}_\epsilon$ *vérifiant* (5.2) *avec une matrice* $A \in \mathbb{R}^{n\times n}_{\min}$ *irréductible. Supposons que les valeurs propres* $\mathcal{L}^1_\epsilon, \ldots, \mathcal{L}^n_\epsilon$ *de* $\mathcal{A}_\epsilon$ *(comptées avec leur multiplicités) ont des asymptotiques de la forme* $\mathcal{L}^i_\epsilon \sim \lambda_i \epsilon^{\Lambda_i}$*. Notons* $\Lambda = (\Lambda_1 \leq \cdots \leq \Lambda_n)$ *la suite des exposants (comptés avec leur multiplicités), et* $\Gamma = (\gamma_1 \leq \cdots \leq \gamma_n)$ *la suite des racines du polynôme caractéristique min-plus de* A*. Alors*

$$\Lambda \prec^{\mathrm{w}} \Gamma \ , \tag{5.3}$$

et pour des valeurs génériques de $a = (a_{ij}) \in \mathbb{C}^{n\times n}$*, on a* $\Lambda = \Gamma$*.*

Ici on a noté $\prec^{\mathrm{w}}$ l'ordre partiel de *dominance faible*, défini pour des vecteurs $u, v \in \mathbb{R}^n_{\min}$ par $u \prec^{\mathrm{w}} v$ si $u_{(1)} \cdots u_{(k)} \geq v_{(1)} \cdots v_{(k)}$ pour $k = 1, \ldots, n$, où pour tout vecteur $u \in \mathbb{R}^n_{\min}$, $u_{(1)} \leq \cdots \leq u_{(n)}$ désignent les coefficients de u ordonnés dans l'ordre croissant.

5.3 Généralisation du théorème de Višik, Ljusternik, Lidskiĭ

L'étude des valeurs propres d'une matrice perturbée de la forme $a + \epsilon b$, où a et b sont des matrices complexes et ϵ est un paramètre, est un problème classique de théorie des perturbations. Une théorie initiée par Višik et Ljusternik, et complétée par Lidskiĭ, permet de déterminer, pour des valeurs génériques de la perturbation b, les asymptotiques au premier ordre $\mathcal{L}_\epsilon \sim \lambda\epsilon^\Lambda$ de toutes les valeurs propres $\mathcal{L}_\epsilon$ de la matrice perturbée. Lorsque a est nilpotente, les exposants Λ sont de la forme $1/k$ pour les tailles k des blocs de Jordan de la matrice a. Les coefficients λ sont les valeurs propres de compléments de Schur d'une matrice extraite particulière de b. Certaines situations non génériques donnent lieu à des cas singuliers, qui ont motivé beaucoup de travaux, notamment de Najman [Naj99], Ma et Edelman [ME98], Moro, Burke et Overton [MBO97]. D'autre part, des résultats d'asymptotiques de vecteurs propres existent dans des cas particuliers.

Considérons comme dans la section précédente des matrices $\mathcal{A}_\epsilon$ admettant un développement limité autour de 0^+ de la forme (5.2), avec $a_{ij} \in \mathbb{C}$ et $A_{ij} \in \mathbb{R}_{\min}$. Dans [S1], nous avons développé dans ce cadre une construction à base de compléments de Schur min-plus et usuels qui généralise la théorie de Višik, Ljusternik, Lidskiĭ, et que nous décrivons rapidement ci-après.

On construit d'abord par récurrence une suite de matrices min-plus A_ℓ pour $1 \leq \ell \leq k$, partant de $A_1 = A$, et d'ensembles d'indices décroissants (on considère ici les matrices comme des éléments de $\mathbb{R}_{\min}^{L\times L}$ où L est un ensemble fini d'indices). Pour tout $\ell \geq 1$, on note $\alpha_\ell = \rho_{\min}(A_\ell)$ et C_ℓ l'ensemble des nœuds critiques de A_ℓ. La récurrence est alors la suivante : tant que $C_1 \cup \cdots \cup C_\ell \neq \{1, \ldots, n\}$,

$$A_{\ell+1} = \mathrm{Schur}(C_\ell, \alpha_\ell, A_\ell) \ ,$$

où pour toute matrice $A \in \mathbb{R}_{\min}^{L\times L}$, toute partition $C \cup N$ de L, et tout $\lambda \in \mathbb{R}_{\min} \setminus \{\mathbb{0}\}$ tel que $\lambda \leq \rho_{\min}(A_{CC})$, le λ-*complément de Schur min-plus* de C dans A est donné par $\mathrm{Schur}(C, \lambda, A) = A_{NN} \oplus A_{NC}(\lambda^{-1}A_{CC})^\star\lambda^{-1}A_{CN}$.

Si A est irréductible, alors toutes les matrices A_ℓ le sont, et $C_\ell \neq \emptyset$, donc l'algorithme s'arrête pour un certain $k \leq n$. On prouve que $\alpha_1 < \cdots < \alpha_k$. Les α_ℓ sont appelés *valeurs critiques* de A, et on définit la *multiplicité* de α_ℓ comme le cardinal $\#C_\ell$ de l'ensemble C_ℓ. On définit alors $\beta = (\beta_1 \leq \cdots \leq \beta_n)$ comme la suite des scalaires $\alpha_1 < \cdots < \alpha_k$ comptés avec leur multiplicité. Si Γ est la suite des racines du polynôme caractéristique min-plus de A, on montre que $\Gamma \prec^{\mathrm{w}} \beta$ [S1, Theorem 4.6].

La construction précédente permet aussi de définir une suite de graphes GC_ℓ, appelés critiques car ils correspondent aux arcs réalisant le minimum dans le calcul de α_ℓ. On caractérise alors l'égalité $\Gamma = \beta$ en termes d'existence d'un recouvrement par des circuits disjoints de ces graphes critiques.

Posons maintenant $\hat{A} = \mathrm{diag}(\beta)^{-1}A$, et notons GC le graphe critique de $\hat{A}$. On construit alors, lorsqu'ils peuvent être définis, les compléments de Schur usuels de la matrice $s^1 := a^{\mathsf{GC}}$, dans l'esprit de la théorie de Višik, Ljusternik, Lidskiĭ :

$$s^\ell = \mathrm{Schur}(C^{\ell-1}, s^1),\ t^\ell = s^\ell_{C_\ell C_\ell},\ \text{pour } C^\ell = C_1 \cup \cdots \cup C_\ell,\ \ell = 2, \ldots, k \ . \tag{5.4}$$

Rappelons que si $a \in \mathbb{R}^{L\times L}$, et $C \cup N$ est une partition de L, le complément de Schur de C dans a est défini par $\mathrm{Schur}(C, a) = a_{NN} - a_{NC}(a_{CC})^{-1}a_{CN}$, lorsque a_{CC} est inversible. On peut aussi calculer s^ℓ par récurrence par $s^\ell = \mathrm{Schur}(C_{\ell-1}, s^{\ell-1})$, si les compléments de Schur sont bien définis. Le résultat suivant est un cas particulier de [S1, Theorem 5.1] (on compte les valeurs propres avec leurs multiplicités) :

Théorème 5.3 ([S1, Theorem 5.1]). *Supposons que les matrices $t^1, \ldots, t^k$ sont inversibles, et notons $\lambda_1^\ell, \ldots, \lambda_{\#C_\ell}^\ell$ les valeurs propres de t^ℓ. Alors pour chaque $\ell = 1, \ldots, k$, $\mathcal{A}_\epsilon$ a exactement $\#C_\ell$ valeurs propres d'ordre ϵ^{α_ℓ}, dont les asymptotiques sont données par*

$$\mathcal{L}_\epsilon^{\ell,j} \sim \lambda_j^\ell \epsilon^{\alpha_\ell}, \quad 1 \leq j \leq \#C_\ell \, . \tag{5.5}$$

Ce résultat est encore vrai si on remplace dans la construction précédente le graphe GC par le graphe de saturation d'un vecteur propre quelconque de $\hat{A}$ par rapport à $\hat{A}$. Dans [S1], on obtient aussi dans certains cas les asymptotiques des vecteurs propres.

Notre construction permet de résoudre des cas singuliers de la théorie de Višik, Ljusternik, Lidskiĭ. Considérons l'exemple dégénéré suivant décrit par Wilkinson et Moro, Burke et Overton :

$$\mathcal{A}_\epsilon = \mathcal{A}_0 + \epsilon b, \text{ pour } \mathcal{A}_0 = \left[\begin{array}{ccc|cc} \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \odot & \cdot & \cdot & \boxdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & 1 \\ \boxdot & \cdot & \cdot & \boxdot & \cdot \end{array}\right], \text{ et } b \in \mathbb{C}^{n \times n} \, . \tag{5.6}$$

Ici les points (entourés ou non) correspondent à des 0. La théorie de Višik, Ljusternik, Lidskiĭ prédit dans ce cas 3 valeurs propres équivalentes à $\lambda\epsilon^{1/3}$ pour les trois racines troisième de b_{31} (le coefficient de b correspondant au cercle), et deux valeurs propres équivalentes à $\mu\epsilon^{1/2}$, pour les deux racines quarrées de l'unique coefficient du complément de Schur de $\{3\}$ dans la matrice $\begin{bmatrix} b_{31} & b_{34} \\ b_{51} & b_{54} \end{bmatrix}$.

Si le coefficient b_{31} est nul aucun des 2 ordres de grandeur $\epsilon^{1/3}$ et $\epsilon^{1/2}$ ne peut avoir lieu. On peut pourtant appliquer le théorème 5.3, avec :

$$a = \left[\begin{array}{ccc|cc} b_{11} & 1 & b_{13} & b_{14} & b_{15} \\ b_{21} & b_{22} & 1 & b_{24} & b_{25} \\ 0 & b_{32} & b_{33} & b_{34} & b_{35} \\ \hline b_{41} & b_{42} & b_{43} & b_{44} & 1 \\ b_{51} & b_{52} & b_{53} & b_{54} & b_{55} \end{array}\right], \text{ et } A = \left[\begin{array}{ccc|cc} 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ \infty & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{array}\right] .$$

On obtient $\alpha_1 = \rho_{\min}(A) = 2/5$ et $C_1 = \{1, 2, 3, 4, 5\}$. Ainsi, $\hat{A} = \alpha_1^{-1} A$ et GC est égal au graphe critique de A, qui est formé exactement du circuit $(1, 2, 3, 4, 5, 1)$. Donc

$$a^{\mathsf{GC}} = \left[\begin{array}{ccc|cc} \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & b_{45} & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & 1 \\ b_{51} & \cdot & \cdot & \cdot & \cdot \end{array}\right] .$$

Le théorème 5.3 établit donc que si $b_{45}b_{51} \neq 0$, les 5 valeurs propres de $\mathcal{A}_\epsilon$ ont chacune une asymptotique de la forme $\mathcal{L}_\epsilon \sim \xi\epsilon^{2/5}$, correspondant à chacune des 5 racines ξ de $\xi^5 = b_{45}b_{51}$.

Il reste des cas dégénérés de matrices A pour lesquelles certaines des matrices t_ℓ ne seront jamais inversibles puisque le graphe critique correspondant ne peut être recouvert par des circuits disjoints. Ces cas peuvent être traités, pour des valeurs génériques de a, par la construction présentée dans la section suivante.

5.4 Perturbation de valeurs propres et problème d'affectation optimal

Afin d'éliminer certains cas de dégénérescence de la section précédente, on considère ici le problème plus général des perturbations de valeurs propres de faisceaux de matrices.

Soit donc un faisceau matriciel $\mathcal{A}_\epsilon = \mathcal{A}_{\epsilon,0} + X\mathcal{A}_{\epsilon,1} + \cdots + X^d\mathcal{A}_{\epsilon,d}$, où X est une indéterminée, et où pour tout $0 \leq k \leq d$, $\mathcal{A}_{\epsilon,k}$ est une matrice $n \times n$ dont les coefficients, $(\mathcal{A}_{\epsilon,k})_{ij}$, sont des fonctions continues à valeurs complexes du paramètre ϵ, vérifiant $(\mathcal{A}_{\epsilon,k})_{ij} = (a_k)_{ij}\epsilon^{(A_k)_{ij}} + o(\epsilon^{(A_k)_{ij}})$ quand ϵ tend vers 0^+, pour $1 \leq i, j \leq n$. On note $a_k = ((a_k)_{ij}) \in \mathbb{C}^{n\times n}$ et $A_k = ((A_k)_{ij}) \in \mathbb{R}_{\min}^{n\times n}$. Les *valeurs propres* $\mathcal{L}_\epsilon$ de $\mathcal{A}_\epsilon$ sont par définition les racines du polynôme (caractéristique) $\det(\mathcal{A}_\epsilon)$, et nous cherchons un équivalent asymptotique de la forme $\mathcal{L}_\epsilon \sim \lambda\epsilon^\Lambda$, pour chaque valeur propre de $\mathcal{A}_\epsilon$. Lorsque $\mathcal{A}_\epsilon = \mathcal{A}_{\epsilon,0} - XI$, on retrouve le problème de la section précédente.

Considérons le *faisceau matriciel min-plus* $A = A_0 \oplus XA_1 \oplus \cdots \oplus X^dA_d$. Les coefficients de A sont donc des polynômes formels à coefficients dans $\mathbb{R}_{\min}$. On appelle *polynôme caractéristique min-plus* du faisceau A, le permanent $P_A = \operatorname{per} A$, qui est un polynôme formel. Notons $\gamma_1, \ldots, \gamma_N$ les racines de P_A, aussi appelées *valeurs propres* du faisceau min-plus A. En généralisant le théorème 5.2, on obtient que pour des valeurs génériques de $a_0, \ldots, a_d$, les valeurs propres du faisceau A sont les exposants des valeurs propres du faisceau $\mathcal{A}$.

Pour calculer les coefficients des valeurs propres, nous construisons un graphe de même nature que le graphe critique GC dans le théorème 5.3. Rappelons que si $B \in \mathbb{R}_{\min}^{n\times n}$, $\operatorname{per} B$ est la valeur d'une affectation optimale dans le graphe associé à B, i.e. le minimum sur toutes les permutations σ de $\{1, \ldots, n\}$ des poids $|\sigma|_B := B_{1\sigma(1)} \cdots B_{n\sigma(n)}$. Si $\operatorname{per} B \neq \mathbb{0}$, on note $\operatorname{Opt}(B)$ le sous-graphe de $\mathsf{G}(B)$ formé des arcs participant à une affectation optimale, i.e. des arcs (i, j) tels qu'il existe une permutation σ vérifiant $j = \sigma(i)$ et $|\sigma|_B = \operatorname{per} B$.

Dans l'énoncé du théorème qui suit, les valeurs propres sont comptées avec leurs multiplicités.

Théorème 5.4 ([J14, Theorem 0.1 ou 1.1])**.** *Soit γ une racine finie de P_A. Pour chaque $0 \leq k \leq d$, notons G_k le sous-graphe de $\operatorname{Opt}(\hat{A}(\gamma))$ formé des arcs (i, j) tels que $\gamma^k(A_k)_{ij} = \hat{A}_{ij}(\gamma)$. Introduisons le faisceau $a^{(\gamma)} := a_0^{G_0} + Xa_1^{G_1} + \cdots + X^d a_d^{G_d}$. Alors, si le faisceau $a^{(\gamma)}$ a m_γ valeurs propres non nulles, $\lambda_1^{(\gamma)}, \ldots, \lambda_{m_\gamma}^{(\gamma)}$, le faisceau $\mathcal{A}_\epsilon$ admet m_γ valeurs propres $\mathcal{L}_\epsilon^{(\gamma),1}, \ldots, \mathcal{L}_\epsilon^{(\gamma),m_\gamma}$ ayant des équivalents respectifs de la forme $\mathcal{L}_\epsilon^{(\gamma),i} \sim \lambda_i^{(\gamma)}\epsilon^\gamma$. En outre, si 0 est une valeur propre de multiplicité m'_γ du faisceau $a^{(\gamma)}$, le faisceau $\mathcal{A}_\epsilon$ admet m'_γ valeurs propres supplémentaires $\mathcal{L}_\epsilon$ telles que $\epsilon^{-\gamma}\mathcal{L}_\epsilon$ converge vers 0, et toutes les autres valeurs propres $\mathcal{L}_\epsilon$ de $\mathcal{A}_\epsilon$ sont telles que le module de $\epsilon^{-\gamma}\mathcal{L}_\epsilon$ tend vers l'infini. Enfin, pour des valeurs génériques des paramètres $(a_k)_{ij}$, m_γ coïncide avec la multiplicité de la racine γ de P_A, et m'_γ coïncide avec la somme des multiplicités des racines de P_A strictement supérieures à γ.*

Dans [J14, Theorem 0.1 ou 1.1], on établit aussi que les graphes G_k peuvent être remplacés par des graphes construits de manière similaire, en remplaçant $\operatorname{Opt}(\hat{A}(\gamma))$ par le graphe de saturation d'une “paire hongroise” (solution du problème dual) du problème d'affectation optimal associé à $\hat{A}(\gamma)$, ce qui conduit à un algorithme efficace de calcul des asymptotiques des valeurs propres décrites dans le théorème 5.4. La preuve de ce théorème repose justement sur un changement d'échelle dépendant d'une telle paire hongroise.

5.5 Perspectives / Travaux en cours

Les résultats présentés ici nécessitent d'être complétés. En effet, le problème du calcul des asymptotiques de vecteurs propres correspondant à des valeurs propres complexes n'est pas complètement résolu : on obtient des résultats de convergence de vecteurs normalisés convenablement, mais les limites peuvent être nulles. Pour résoudre ce problème, et celui du calcul des valeurs propres dans les cas dégénérés, il faudrait développer un algorithme de Puiseux matriciel, en s'inspirant par exemple de l'algorithme introduit par Murota [Mur90].

Les preuves des théorèmes 5.3 et 5.4 font appel à des changement de variable diagonaux, qui permettent de changer d'échelle et de se concentrer autour d'un ordre de grandeur de valeur propre. Ceci suggère que l'on pourrait utiliser le même genre de changement d'échelle pour améliorer la précision du calcul numérique de valeurs propres de matrices ayant des coefficients d'ordre de grandeur variables. On pourrait aussi l'utiliser dans le problème relié du calcul de pseudo-spectres, c'est-à-dire des ensembles de "valeurs propres approchées". Cette direction de recherche a commencé à être explorée au cours de 3 stages de M2 (en 05, 06 et 07) que j'ai co-encadré avec Stéphane Gaubert (INRIA Rocquencourt), et pour celui de 05 avec Papa Momar Ndiaye (Raise Partner Sas, Grenoble).

D'autres directions de recherche possibles n'ont pas encore été étudiées. On peut par exemple penser à la généralisation au cas de la dimension infinie des résultats précédents sur les perturbations de valeurs propres (perturbations d'opérateurs ou perturbations singulières de systèmes dynamiques à la Freidlin-Wentzell), ainsi que le problème de la limite quand la taille des matrices tend vers l'infini (comme on le rencontre en mécanique statistique).

Finalement, les développements récents en géométrie algébrique tropicale [GKZ94, Vir01, Mik04, SS04] suggèrent de nouvelles directions de recherche : calcul formel, équations polynomiales, variétés algébriques max-plus.

Deuxième partie

Applications contractantes au sens large

Chapitre 6

Applications monotones homogènes en programmation dynamique

Nous étudions ici quelques problèmes relatifs aux applications monotones additivement homogènes (ou sous-homogènes) qui sont soulevés par l'étude de problèmes de contrôle optimal ou de jeux à somme nulle : structure de l'espace propre, asymptotiques des itérées, longueur des orbites périodiques. Les conditions employées pour résoudre ces problèmes sont selon les cas la convexité, la semi-différentiabilité, la continuité.

6.1 Un théorème spectral pour le contrôle stochastique

Les travaux présentés ici ont été obtenus en collaboration avec Stéphane Gaubert (INRIA Rocquencourt). Ils ont fait l'objet des articles publiés [C14, J12].

On considère ici une application $f : \mathbb{R}^n \to \mathbb{R}^n$ monotone, additivement homogène et convexe, i.e. qui vérifie les conditions (M) et (AH) de l'introduction pour tous $x, y \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$, ainsi que la convexité des applications coordonnées $f_i : \mathbb{R}^n \to \mathbb{R}$. On suppose que cette application a au moins un vecteur propre (additif), i.e. qu'il existe $\lambda \in \mathbb{R}$ et $v \in \mathbb{R}^n$ tels que $f(v) = \lambda + v$. Comme f est contractante au sens large pour la norme infinie, la valeur propre λ est nécessairement unique. On note $\mathscr{E}(f)$ l'*espace propre* de f, i.e. l'ensemble des solutions $v \in \mathbb{R}^n$ à l'équation $f(v) = \lambda + v$. Dans [J12], nous avons montré pour ce type d'applications le théorème spectral 6.1 ci-dessous qui généralise le théorème spectral max-plus. Notons que contrairement au cas des applications max-plus linéaires, nous ne considérons pas des vecteurs propres prenant la valeur $-\infty$, même si une extension continue de f à $(\mathbb{R} \cup \{-\infty\})^n$ existe toujours [BNS03].

Pour établir le théorème spectral, nous avons besoin comme dans le cas max-plus linéaire de la notion de graphe critique. Pour le définir, on note, pour tout $x \in \mathbb{R}^n$, $\partial f(x) = \{P \in \mathbb{R}^{n\times n} \mid f(y) - f(x) \geq P(y - x),\ \forall y \in \mathbb{R}^n\}$, le *sous-différentiel* de f, qui n'est autre que le produit cartésien des sous-différentiels usuels des fonctions coordonnées. On montre facilement que $\partial f(x)$ n'est formé que de matrices stochastiques. Si v est vecteur propre quelconque de f, on définit le *graphe critique* de f, noté $\mathsf{GC}(f)$, comme l'union des graphes finaux des matrices stochastiques $P \in \partial f(v)$, où pour toute matrice stochastique P, le graphe final de P est la restriction du graphe de P aux nœuds des classes finales. On montre que le graphe $\mathsf{GC}(f)$ ainsi défini est indépendant du choix du vecteur propre v. Les nœuds de $\mathsf{GC}(f)$ sont appelés *nœuds critiques*. Les ensembles de nœuds des composantes fortement connexes de $\mathsf{GC}(f)$ sont appelés *classes critiques*. On définit la cyclicité du graphe critique de la même façon que dans le cas max-plus linéaire (cf. le paragraphe juste avant le théorème 2.2) et on le note $\sigma(f)$.

On dit d'une application $g : U \subset \mathbb{R}^n \to V \subset \mathbb{R}^p$ qu'elle est un *isomorphisme* monotone homogène si g et g^{-1} sont toutes deux monotones homogènes. Rappelons qu'un inf-sous-semi-treillis d'un inf-semi-treillis $\mathscr{T}$ est un sous-ensemble $\mathscr{S}$ de $\mathscr{T}$ tel que, pour tous $s, t \in \mathscr{S}$, l'infimum de s et t dans $\mathscr{T}$ (qui existe par définition d'un inf-semi-treillis) est un élément de $\mathscr{S}$.

Théorème 6.1 ([J12, Theorems 3.4, 6.6, Corollaries 3.6, 5.7])**.** *Considérons une application monotone homogène convexe $f : \mathbb{R}^n \to \mathbb{R}^n$ qui a un vecteur propre : $\mathscr{E}(f) \neq \emptyset$. Notons λ l'unique valeur propre de f, C l'ensemble de ses nœuds critiques, et $\sigma = \sigma(f)$. Alors,*

1. *la restriction $\mathsf{r} : \mathbb{R}^n \to \mathbb{R}^C, x \mapsto (x_i)_{i \in C}$, est un isomorphisme monotone homogène de $\mathscr{E}(f)$ dans son image $\mathscr{E}^{\mathrm{c}}(f)$;*
2. *$\mathscr{E}^{\mathrm{c}}(f)$ est un inf-sous-semi-treillis de $(\mathbb{R}^C, \leq)$;*
3. *$\mathscr{E}^{\mathrm{c}}(f)$ est un convexe de dimension inférieure ou égale au nombre de classes critiques de f, avec égalité si f est affine par morceaux;*
4. *pour tout $x \in \mathbb{R}^n$, $f^{kc}(x) - kc\lambda$ a une limite quand $k \to \infty$.*

En particulier lorsque f n'a qu'une classe critique, le vecteur propre de f est unique à une constante additive près.

Le cas d'applications f affines par morceaux correspond aux problèmes de contrôle optimal stochastique avec un ensemble de contrôles U fini. Dans ce cas particulier, les différentes assertions du théorème 6.1 peuvent être obtenues en rassemblant les résultats de Lanery [Lan67], Romanovsky [Rom73], et Schweitzer et Federgruen [SF78, SF77].

Une idée importante dans la preuve du théorème 6.1 est le principe du maximum pour les matrices stochastiques : si P est une matrice stochastique et si $Pv \leq v$ alors d'une part, sur chaque classe finale de P, $Pv = v$ et donc v est constant, d'autre part le minimum de v sur $\{1, \ldots, n\}$ est atteint pour un indice i appartenant à une classe finale de P. Cette idée peut ainsi se transposer dans le cadre des équations aux dérivées partielles en utilisant la notion de solution de viscosité. C'est ce qui a été fait dans [J8] et dans le travail de stage de M2 et de début de thèse de Benoît David [S3]. Les autres ingrédients sont les dérivées directionnelles et pour la cyclicité le théorème de Nussbaum [Nus90] et Sine [Sin90] sur la convergence des orbites de fonctions contractantes au sens large pour la norme infinie.

6.2 Applications semi-différentiables monotones additivement homogènes et jeux à somme nulle

Les travaux présentés ici brièvement font partie d'un travail en cours, en collaboration avec Stéphane Gaubert et Roger Nussbaum (Rutgers University, USA). Ils font l'objet de deux articles en préparation [P2, P5].

Le but ici est de généraliser le théorème spectral de la section précédente au cas d'applications monotones additivement homogènes non nécessairement convexes. Dans ce cas, la dimension de l'espace propre ne peut pas être définie : celui-ci n'est pas convexe, et la dimension "locale" peut varier. Par contre nous montrons que le théorème d'unicité se généralise d'une certaine manière.

On se place plus généralement dans le cadre des applications monotones positivement homogènes de l'intérieur d'un cône d'un espace de Banach dans lui-même. On dit qu'une application f d'un espace normé E dans un autre, est semi-différentiable si l'on peut écrire $f(x+h) = f(x) + f'_x(h) + o(\|h\|)$, où $f'_x(h)$ est une application continue positivement homogène. Cette notion, qui a été introduite par Penot, est bien adaptée aux opérateurs de programmation dynamique associés à des problèmes de contrôle optimal ou de jeux : lorsque les espaces d'actions sont finis, f, qui est affine par morceaux, est non-différentiable, mais elle est semi-différentiable, au moins lorsque l'espace d'état est lui même fini.

Dans le cas différentiable, des résultats antérieurs de Nussbaum [Nus88] montrent que l'unicité du vecteur propre de f, à une normalisation près, peut être contrôlée à l'aide de résultats d'unicité pour la dérivée f'_v, évaluée en un vecteur propre quelconque v. Nous montrons qu'il en est de même lorsque f est semi-différentiable. On montre aussi des résultats d'unicité pour les points fixes d'applications monotones sous-homogènes, ainsi que la convergence géométrique des itérées de telles applications.

On retrouve ainsi la spécialisation au cas d'une seule classe critique du résultat de la section précédente. L'application de ces résultats au cas d'applications monotones additivement homogènes de $\mathbb{R}^n$ dans lui-même fournit des résultats d'unicité pour la solution d'équations de la programmation dynamique stationnaires, ou ergodiques, associées à certains problèmes de jeux à somme nulle, ainsi que des résultats de convergence géométrique pour l'itération sur les valeurs, avec un contrôle explicite du taux de convergence.

6.3 Itérées d'applications monotones sous-homogènes

Les travaux présentés ici ont été obtenus en collaboration avec Stéphane Gaubert (INRIA Rocquencourt), Bas Lemmens (Warwick University, UK) et Roger Nussbaum (Rutgers University, USA). Ils ont fait l'objet de l'articles publié [J15].

Étant donné une application monotone additivement sous-homogène de $\mathbb{R}^n$ dans lui-même, sa conjuguée par l'application exponentielle terme à terme est une application monotone positivement sous-homogène de l'intérieur $(\mathbb{R}_+^*)^n$ du cône positif dans lui-même. C'est donc aussi une application contractante au sens large pour la métrique de Thompson (qui est la conjuguée de la norme infinie). Or pour les applications de l'intérieur d'un cône polyhédral K dans lui-même, qui sont contractantes au sens large pour la métrique de Thompson de K, il est connu que les orbites bornées qui n'approchent pas la frontière du cône convergent vers une orbite périodique. De plus, dans le cas où K est le cône positif, les longueurs des orbites sont bornées par une constante ne dépendant que de n, et des bornes ont été conjecturées ou prouvées selon que l'on se restreint ou non à une sous-classe d'applications. Par exemple, pour les applications monotones homogènes, Gunawardena et Sparrow ont conjecturé que les longueurs des orbites étaient bornées par $\binom{n}{\lfloor n/2 \rfloor}$, ce qui a été prouvé par Lemmens et Scheutzow [LS05] (la borne est optimale).

D'autre part, comme on l'a déjà rappelé plus haut (dans le cadre additif), toute application monotone sous-homogène de $(\mathbb{R}_+^*)^n$ dans $\mathbb{R}_+^n$ admet une extension continue en une application monotone sous-homogène de $\mathbb{R}_+^n$ dans lui-même [BNS03]. On se pose donc ici la question de la convergence des orbites bornées (ou de manière équivalente bornées supérieurement) d'applications monotones sous-homogènes continues de $\mathbb{R}_+^n$ dans lui-même ou plus généralement d'un cône polyhédral K fermé dans lui-même. Pour un tel cône, on appelle face, une face de dimension $\dim K - 1$. On a le résultat général de convergence suivant.

Théorème 6.2 ([J15, Theorem 2.1])**.** *Soit K un cône polyhédral avec N faces. Si $f\colon K \to K$ est une application monotone sous-homogène continue, et si $x \in K$ a son orbite bornée, alors il existe* *un point ξ périodique pour f de période p tel que* $\lim_{k\to\infty} f^{kp}(x) = \xi$. *De plus* $p \leq \beta_N$, *où*

$$\beta_N = \max_{q+r+s=N} \frac{N!}{q!r!s!} = \frac{N!}{\lfloor \frac{N}{3} \rfloor ! \lfloor \frac{N+1}{3} \rfloor ! \lfloor \frac{N+2}{3} \rfloor !}. \tag{6.1}$$

On montre de plus que dans le cas où K est le cône positif la borne précédente est asymptotiquement optimale [J15, Theorem 2.2]. Cette borne est à comparer avec celle de Lemmens et Scheutzow [LS05] pour les longueurs des orbites contenues dans l'*intérieur* du cône positif :

$$\binom{N}{\lfloor N/2 \rfloor} := \frac{N!}{\lfloor \frac{N}{2} \rfloor ! \lfloor \frac{N+1}{2} \rfloor !},$$

qui est d'ailleurs un des ingrédients de la preuve.

6.4 Perspectives / Travaux en cours

En plus du travail en cours présenté dans la section 6.2, un travail en cours, en collaboration avec Stéphane Gaubert (INRIA Rocquencourt) et Bas Lemmens (Warwick University, UK), traite de la généralisation du théorème spectral convexe 6.1 au cas d'applications monotones convexes qui sont contractantes au sens large pour une norme polyhédrale, ou plus généralement qui admettent un point fixe stable.

D'autres directions de recherche restent encore à développer, en particulier dans le cadre de la dimension infinie. Par exemple, les problèmes de contrôle optimal stochastique à temps discret et espace dénombrable contiennent en particulier les problèmes de contrôle optimal de réseaux de files d'attente. L'étude de tels problèmes pourrait se faire via une construction englobant à la fois les techniques présentées dans la section 6.1 et celles du chapitre 2, comme la tension ou la frontière de Martin max-plus.

Les problèmes de contrôle stochastique à temps et espace continu peuvent se traiter soit via les semi-groupes à temps continu d'opérateurs monotones homogènes $(S^t)_{t\geq 0}$, soit via les équations aux dérivées partielles d'Hamilton-Jacobi-Bellman. En particulier, le problème ergodique conduit à l'étude des vecteurs propres v du semi-groupe, $S^t v = \lambda t + v$, ou à l'étude des solutions v et λ de l'équation d'Hamilton-Jacobi-Bellman ergodique, $\lambda - H(x, \frac{\partial v}{\partial x}, \frac{\partial^2 v}{\partial x^2}) = 0$. Le développement de résultats de structure dans ce cadre pour les vecteurs propres a encore été peu abordé, que ce soit avec des techniques de semi-groupes ou d'équations aux dérivées partielles, alors que pour le cas déterministe la théorie des solutions KAM faible de Fathi a déjà fait l'objet de nombreux développements. Pour ce qui est du contrôle stochastique, on peut citer les résultats d'unicité obtenus, dans le cas d'un bruit constant et avec des techniques variationnelles, par Bensoussan [Ben88]. Dans [J8], nous avons obtenu, en collaboration avec A. Sulem (INRIA Rocquencourt) et M. Taksar (University of Missouri), un résultat d'unicité dans le cas particulier d'un problème de gestion de portefeuille en temps long, en utilisant la notion de solution de viscosité à la Crandall-Ishii-Lions de l'inéquation variationnelle correspondante, et en s'inspirant de la technique de [J12]. Un résultat récent de Barles et Da Lio [BDL05] donne aussi un résultat d'unicité dans le cas d'une équation fortement elliptique, en utilisant des techniques de solutions de viscosité. Il reste encore beaucoup à faire. Le travail de stage de M2 et de début de thèse de Benoît David porte sur ce sujet. Une première généralisation à certains cas particuliers de problèmes de contrôle stochastique, admettant un nombre fini de points d'équilibres stables, a déjà été établie [S3]. On pourrait aussi envisager des résultats d'unicité pour des problèmes de jeux à somme nulle.

Chapitre 7

Applications monotones contractantes au sens large intervenant dans d'autres domaines

Nous avons regroupé ici des travaux qui ne concernent pas la théorie du contrôle optimal, mais qui utilisent partiellement les techniques de fonctions monotones contractantes au sens large.

7.1 Étude de quelques systèmes à retard oscillants

Les travaux présentés ici ont été obtenus en collaboration avec Pierre-Alexandre Bliman (INRIA Rocquencourt), Michel Sorine (INRIA Rocquencourt) et Sophie Bismuth (Postdoc à l'INRIA Rocquencourt au moment des travaux). Ils ont fait l'objet des articles publiés [C10, C11, J7, J9, J10].

Vue d'ensemble des résultats

La modélisation et/ou la commande de certains systèmes physiques conduit à des équations différentielles à retard comprenant de fortes non–linéarités, voire des discontinuités, de telle sorte que le système a un comportement oscillatoire, par exemple périodique.

Ce phénomène apparaît par exemple dans la commande moteur, où la sonde de richesse fournit une information binaire et retardée de la richesse des gaz d'échappement à la sortie d'un moteur thermique. La richesse x, normalisée par rapport à une valeur stœchiométrique, peut être modélisée par le système

$$\dot{x}(t) + x(t) = u(t), \quad y(t) = \mathrm{sgn}(x(t-1)) \ ,$$

où x est l'état, u est le contrôle et y est la sortie. Ici, pour simplifier les notations, on a normalisé la constante de temps et le retard à 1. Dans [J10], nous avons étudié les commandes proportionnelle et proportionnelle-intégrale. Celles-ci conduisent à l'étude du système :

$$\begin{cases} \dot{x}(t) = -(k_I\xi(t) + k_P\mathrm{sgn}(t, x(t-1)) + x(t)) \ , \\ \dot{\xi}(t) = \mathrm{sgn}(t, x(t-1)) \ , \end{cases} \tag{7.1}$$

où $k_I = 0$ et $k_P > 0$ dans le cas de la commande proportionnelle, et $k_I > 0$ dans le cas de la commande proportionnelle-intégrale. Ici les fonctions x et ξ sont absolument continues, et on a remplacé la fonction signe par une fonction dépendant éventuellement du temps mais prenant uniquement (ou presque sûrement) les valeurs ± 1. En fait il suffit que la fonction signe ne prenne jamais la valeur 0, afin d'éliminer la solution $x \equiv 0$ qui serait instable.

Dans le cas d'un retard nul, ces commandes permettent de ramener les trajectoires à 0, malgré la pauvreté de la sortie. Dans le cas d'un retard non nul comme ici, les commandes conduisent à des oscillations, dont la moyenne est nulle.

Dans [J10, Theorem 1], on montre que, pour les deux types de commande, (7.1) a une infinité dénombrable $(x_n, y_n)_{n\in\mathbb{N}}$ de solutions périodiques à 2 phases, i.e. qui changent une seule fois de signe dans un intervalle $[t, t+T_n]$ bien choisi, où T_n est la période, et que celles-ci sont symétriques (i.e. telles que $x_n(t+T_n/2) = -x_n(t)$ et $y_n(t+T_n/2) = -y_n(t)$). Les périodes T_n sont telles que $1/(n+1/2) < T_n < 1/n$. Ainsi, parmi ces solutions, une seule est lentement oscillante (i.e. telle que la distance entre deux de ses zéros est toujours strictement supérieure au retard). On montre que cette dernière est localement orbitalement asymptotiquement stable, que l'on peut identifier les paramètres du problème au moyen de sa période et de son amplitude, et qu'inversement on peut contrôler sa période et son amplitude au moyen des paramètres de commande k_P et k_I. On montre aussi un résultat de type rejet de perturbations.

L'étude de la commande proportionnelle, déjà traitée en partie dans les travaux de L. Fridman, Shustin et E. Fridman [FSF96, Shu95], a été menée jusqu'au bout. On montre en effet que toute solution converge en temps fini vers une solution périodique [J10, J7], et que les solutions périodiques non lentement oscillantes sont instables [J10, Theorem 2]. Nous détaillons plus bas ces résultats, car les preuves se ramènent à l'étude des itérées d'applications contractantes au sens large pour la norme ℓ^1.

Dans le prolongement de ces travaux, nous avons étudié, dans [J9], le système simple $\dot{x} = F(x) - h(x(t-1))$, où h est une non-linéarité constante par morceaux, antisymétrique, non monotone prenant au plus 4 valeurs, et F est bornée par le minimum de la valeur absolue de h. On prouve l'instabilité de toutes les solutions rapidement oscillantes, au moyen d'une étude locale (valeurs propres).

Résultats à base de fonctions contractantes au sens large

Nous considérons ici le cas de la commande proportionnelle. En considérant la fonction x/k_P, le système (7.1) se ramène à l'équation

$$\dot{x}(t) = -(\mathrm{sgn}(t, x(t-1)) + x(t)) \; . \tag{7.2}$$

On définit $V(t)$ comme le cardinal du nombre de zéros avec changement de signe dans l'intervalle de longueur finissant au premier zéro suivant t :

$$Z = \{t \geq 0 \mid x(t) = 0 \text{ et } \forall \varepsilon > 0,\ \exists t' \in [t-\varepsilon, t),\ t'' \in (t, t+\varepsilon],\ x(t')x(t'') < 0\} \; , \tag{7.3}$$

$$V(t) = \# Z \cap [t' - h, t') \text{ où } t' = \inf\, [t, +\infty) \cap Z \; . \tag{7.4}$$

Alors $V(t) \equiv 0$ pour t grand lorsque x est lentement oscillante à partir d'un certain temps. De plus, pour tout n, la solution périodique x_n de la section précédente vérifie $V(t) \equiv 2n$. En général $V(t) \in \mathbb{N} \cup \{+\infty\}$ pour tout $t \geq 0$.

La fonction V joue le rôle d'une fonction de Lyapounov :

Théorème 7.1 ([J10, Theorem 2]). *Pour toute solution au problème de Cauchy associé à* (7.2)*, $V(t)$ est décroissante et prend des valeurs paires (ou infinies).*

Si $2n = \lim_{t\to+\infty} V(t) < +\infty$ *et* $\underline{t}$ *est le plus petit* t *tel que* $V(t) = 2n$*, alors il existe* $\varphi \geq 0$ *tel que*

$$t \geq \underline{t} \Rightarrow x(t) = x_n(t+\varphi) \ .$$

La solution périodique lentement oscillante x_0 *est localement orbitalement asymptotiquement stable. Les solutions non lentement oscillantes* x_n*,* $n \geq 1$*, sont instables.*

Ce résultat prouve en particulier que les solutions x_n sont les seules solutions périodiques de (7.2).

La preuve de l'instabilité des solutions non lentement oscillantes repose sur l'assertion suivante : si $V(t) = 2n$ pour tout $t \geq t_0$ (ce qui serait le cas si on partait d'une condition initiale proche de x_n et si x_n était stable), alors x est au déphasage près, la solution périodique x_n. Pour prouver cela on construit une application Φ du simplexe $\Sigma_n := \{s \in \mathbb{R}_+^{2n+1} \mid \sum_{i=0,\ldots,2n} s_i = 1\}$ dans lui-même qui représente l'évolution en temps rétrograde du système. Si on note t_k la suite croissante des éléments de Z, et $b^k \in \Sigma_n$ la suite des longueurs des intervalles de $[t_k - 1, t_k)$ où x est de signe constant : $b_i^k = t_{k-2n+i} - t_{k-2n+i-1}$ pour $i = 1, \ldots, 2n$, $b_0^k = 1 - (b_1^k + \cdots + b_{2n}^k)$, et l'application Φ associe b^{k-1} à b^k. Cette application est toujours bien définie et envoie l'intérieur du simplexe dans lui-même, alors que son application inverse n'est bien définie que pour des b tels que $V(t)$ reste constant. On voit facilement que Φ est monotone, préserve le simplexe, et que sa différentielle est de norme ℓ^1 égale à 1. Ainsi, Φ est contractante au sens large pour la norme ℓ^1. Dans [J10], on montre par des arguments de signe et en utilisant les différentielles, que ϕ^N est contractante strictement pour la norme ℓ^1 pour un certain N dépendant seulement de n. Ceci implique que Φ a un point fixe unique, lequel correspond à la solution périodique x_n, et que si $b^k = \phi(b^{k+1})$ pour une suite $(b^k)_{k\geq 0}$ dans Σ_n, donc bornée, alors b^0 est nécessairement un point fixe.

L'énoncé de [J10, Theorem 2] contient en plus la propriété suivante :

Théorème 7.2 ([J7, Theorem 2])**.** *Pour toute solution au problème de Cauchy associé à* (7.2)*,* $V(t)$ *est fini à partir d'un certain temps* t_{x_0} *tel que*

$$t_{x_0} \leq C_\varepsilon \left(1 + \delta^{-2-\varepsilon}\right) \tag{7.5}$$

où δ *est le supremum des longueurs des intervalles de* $(0,1)$ *où* x *est de signe constant.*

Ce résultat a été prouvé dans [J7] pour l'équation différentielle plus générale :

$$\dot{x}(t) = -\mathrm{sgn}(t, x(t-1)) + F(x(t)) \ , \tag{7.6}$$

où $F : \mathbb{R} \to \mathbb{R}$ est mesurable au sens de Lebesgue, telle que $\|F\|_{L^\infty(\mathbb{R})} < 1$ (noter que toute solution de (7.2) vérifie $\sup_{t\geq t_0} |x(t)| < 1$ pour t_0 assez grand).

La preuve, inspirée de [Shu95], repose aussi sur un argument similaire à celui de l'instabilité des solutions périodiques non lentement oscillantes. Si $V(t) \equiv +\infty$, alors Z a des points d'accumulation, et si I est un intervalle de longueur 1 dont les bornes sont des points d'accumulation, on lui associe une quantité $\mu(I)$ qui est la somme sur tous les intervalles maximaux $J \subset I$ où x est de signe constant, d'une quantité $\phi(\mathrm{sgn}(J), \mathrm{mes}\,(J))$ dépendant du signe $\mathrm{sgn}(J)$ de x sur l'intervalle J et de la longueur de l'intervalle J. On montre pour une grande quantité de fonctions ϕ convexes par rapport à la deuxième variable, que $\mu(I) \leq \mu(1+I)$, et on montre une contraction stricte pour une classe restreinte de fonctions ϕ, telles que $\mu(I)$ est équivalent à la somme des $\mathrm{mes}\,(J)^{1+\varepsilon}$ pour les intervalles $J \subset I$ définis ci-dessus, conduisant à l'estimation (7.5). Le cas limite de la norme ℓ^1, correspondant à $\epsilon = 0$ ne peut donner d'estimation puisque celle-ci est conservée.

D'autres preuves de ce résultat dans des cas particuliers, utilisant des techniques d'applications monotones contractantes au sens large pour la norme ℓ^1 sont apparues dans [NS01].

7.2 Un modèle de parcours auto-validant du Web

Les travaux présentés ici ont été obtenus en collaboration avec Stéphane Gaubert (INRIA Rocquencourt), et Laure Ninove (Doctorante de l'UCL, Belgique). Ils ont fait l'objet de l'article publié [B9] et de la pré-publication [S4].

Le "PageRank" le plus couramment utilisé est un ordre des pages Web calculé à partir du graphe du Web, lequel a pour noeuds les pages Web et pour arcs les hyperliens : on suppose qu'un visiteur d'une page Web choisit la page suivante qu'il va visiter de manière uniforme parmi les pages pointées par la page qu'il est en train de visiter, ainsi le parcours d'un visiteur du Web est une marche aléatoire sur le graphe du Web et on ordonne les pages Web par la valeur de la mesure invariante $\mathbf{r}$ de cette marche aléatoire. Si on note $C \in \mathbb{R}_+^{n\times n}$ la matrice d'adjacence du graphe du Web (les pages sont numérotés de 1 à n, et $C_{ij} = 1$ si il y a un hyperlien de la page i à la page j, et $C_{ij} = 0$ sinon), alors la probabilité de transition de la page i à la page j est égale à :

$$M_{ij} = \frac{C_{ij}}{\sum_k C_{ik}} ,$$

et le "PageRank" $\mathbf{r}$ est le vecteur de probabilité vérifiant $\mathbf{r} = \mathbf{r}M$ (qui est unique lorsque le graphe du Web est fortement connexe).

L'hypothèse qu'un visiteur du Web puisse parcourir les pages suivantes pointées de manière uniforme ne semblant pas réaliste, nous avons proposé un modèle tenant compte du fait qu'un visiteur du Web peut avoir une idée a priori de la valeur des pages, et ainsi favoriser, dans ses choix de page suivante, les pages "réputées". Nous avons ainsi proposé un modèle dans lequel la réputation tient compte du PageRank, la probabilité d'aller de i à j étant proportionnelle à $C_{ij}e^{\mathbf{r}(s)_j/T}$. Ici $T > 0$ est un paramètre fixé qu'on appelle *température*, comme en mécanique statistique. La matrice des probabilités de transition dépend de la valeur courante $\mathbf{r}(s)$ du PageRank : c'est la matrice $M_T(\mathbf{r}(s))$, telle que pour tout $\mathbf{x}$,

$$M_T(\mathbf{x})_{ij} = \frac{C_{ij}e^{\mathbf{x}_j/T}}{\sum_k C_{ik}e^{\mathbf{x}_k/T}} . \tag{7.7}$$

Lorsque T tend vers l'infini, on retrouve la matrice M usuelle. Lorsque T tend vers 0^+, le visiteur ne visite que les pages de rang maximal parmi celles qui lui sont accessibles. Supposons qu'à intervalles réguliers (par exemple une fois par semaine), le Pagerank soit modifié pour tenir compte de l'état du système. Alors à l'étape s, le nouveau PageRank $\mathbf{r}(s+1)$ sera la mesure invariante pour la matrice de transition $M_T(\mathbf{r}(s))$. Ainsi si on note $\mathbf{u}_T(\mathbf{x})$ la mesure invariante de $M_T(\mathbf{x})$ (qui est unique lorsque C est irréductible), on obtient $\mathbf{r}(s+1) = \mathbf{u}_T(\mathbf{r}(s))$. On appelle *$T$-PageRank* la limite de $\mathbf{r}(s)$ lorsque s tend vers l'infini, si elle existe.

Dans [B9, S4], on montre que lorsque T est grand ($T \geq n$), l'application $\mathbf{u}_T$ a un unique point fixe à coefficients strictement positifs. Si de plus C est primitive, toutes les orbites convergent vers ce point fixe. La preuve de ce résultat est basé sur un résultat de Nussbaum [Nus88] sur les fonctions monotones sous-homogènes. En effet en utilisant la formule des réseaux fluviaux (formule de Tutte), on obtient pour $\mathbf{u}_T(\mathbf{x})$ une formule qui à une normalisation près est une application monotone. Elle est de plus sous-homogène si $T \geq n$, ce qui conduit au résultat.

Par contre, si la température est petite, il peut exister plusieurs T-PageRanks, et la limite de la suite $\mathbf{r}(s)$ dépendra de la croyance initiale $\mathbf{r}(0)$. Ainsi, des effets autovalidants apparaissent.

D'autres résultats ont été établis, par exemple lorsque C est primitive et T est grand, alors le PageRank est l'unique limite de la suite $\tilde{\mathbf{r}}(s)$ définie par la récurrence :

$$\tilde{\mathbf{r}}(s+1) = \tilde{\mathbf{r}}(s) M(\tilde{\mathbf{r}}(s)) . \tag{7.8}$$

7.3 Perspectives / Travaux en cours

Dans l'étude des systèmes à retards faite dans [J9], les résultats d'instabilité des solutions rapidement oscillantes ont été prouvés au moyen d'une étude locale (valeurs propres). Pourtant, on pourrait envisager de traiter ce problème et ses généralisations au moyen d'applications contractantes au sens large, comme cela a été le cas dans les travaux présentés dans la section 7.1. Par ailleurs, Nussbaum a fait apparaître le lien entre vecteurs propres d'applications max-plus linéaires de dimension infinie et solutions périodiques d'équations différentielles à retard dépendant de l'état [MPN03].

Ces exemples et ceux de la section 7.1 montrent que l'étude (existence, stabilité) des solutions périodiques de certaines équations à retard peut se ramener à l'étude des points fixes d'applications monotones contractantes au sens large. On pourrait chercher d'autres exemples de ce phénomène, par exemple dans le cas de la commande saturée ou proportionnelle-intégrale, afin d'illustrer et d'enrichir les résultats sur les applications contractantes au sens large.

Dans l'étude des pages du Web, la non-linéarité de la dynamique considérée provient du caractère non stationnaire de la chaîne de Markov sous-jacente. On pourrait chercher d'autres exemples où ce genre de chaînes de Markov apparaissent.

Finalement, signalons que certains modèles de dynamique des populations font intervenir des systèmes dynamiques $\dot{x} = f(x)$ non-linéaires, l'application f jouissant de certaines propriétés de monotonie. Les résultats classiques (à la Smith, ou Hirsch, voir par exemple [Smi95]) développés dans ce cadre n'ont pas été rapprochés des résultats de nature plus combinatoire obtenus dans la littérature pour le cas des systèmes dynamiques monotones à temps discret (cf. le chapitre 6). Un travail d'unification devrait être à terme particulièrement fructueux.

Troisième partie

Équations d'Hamilton-Jacobi-Bellman

Chapitre 8

Analyse numérique des équations de la programmation dynamique

Ce chapitre est consacré essentiellement à la résolution numérique des équations aux dérivées partielles de type Hamilton-Jacobi-Bellman associées aux problèmes de contrôle optimal stochastique. Contrairement aux méthodes présentées au chapitre 4 pour le cas déterministe, nous supposons ici donnée une discrétisation "classique" (par exemple par différences finies) de l'équation aux dérivées partielles, telle que l'équation discrétisée s'interprète comme l'équation de la programmation dynamique d'un problème de contrôle stochastique à temps discret et espace d'état fini. Nous cherchons alors à développer des algorithmes rapides pour la résolution de l'équation discrète.

8.1 Méthodes multigrilles pour les problèmes de contrôle stochastique actualisés

Les travaux présentés ici ont été publiés dans la thèse [T1] et les articles [B1, B2]. Voir aussi l'article [C2] en collaboration avec Jean-Pierre Quadrat (INRIA Rocquencourt) et Jean-Philippe Chancelier (CERMICS-ENPC).

Considérons le problème de contrôle de diffusion arrêté et actualisé :

$$v(x) = \inf_{u(.)} \mathbb{E}\left[\int_0^\tau e^{-\lambda t} c(\mathbf{u}_t, \mathbf{x}_t)\, dt + e^{-\lambda\tau}\phi(\mathbf{x}_\tau) \mid \mathbf{x}_0 = x\right] ,$$

où l'état $\mathbf{x}$ est un processus de diffusion vérifiant l'équation différentielle stochastique dans $\mathbb{R}^d$:

$$d\mathbf{x}_t = g(\mathbf{u}_t, \mathbf{x}_t)\, dt + \sigma(\mathbf{x}_t)\, dW_t$$

pour un processus de Wiener W_t dans $\mathbb{R}^d$, où $\lambda > 0$ est un taux d'actualisation, et τ est le temps de sortie de $\mathbf{x}$ d'un ouvert $\Omega \subset \mathbb{R}^d$.

La fonction valeur v est alors solution (sous certaines conditions de régularité) de l'équation d'Hamilton-Jacobi-Bellman :

$$\inf_{u\in U}\{[A(u)\, v + c(u)](x)\} = 0 \quad x \in \Omega \tag{8.1a}$$

$$v(x) = \phi(x) \quad x \in \partial\Omega \tag{8.1b}$$

où $c(u) = c(u, \cdot)$ et

$$A(u)\, v(x) = \sum_{ij} a_{ij}(x) \frac{\partial^2 v}{\partial x_i \partial x_j}(x) + \sum_j g_j(u,x) \frac{\partial v}{\partial x_j}(x) - \lambda v(x) \quad \text{avec } a = \frac{1}{2}\sigma\sigma^T \ . \tag{8.2}$$

On peut aussi réécrire (8.1a) sous forme vectorielle :

$$\inf_{\mathbf{u}\in\mathcal{U}} \{A(\mathbf{u})\, v + c(\mathbf{u})\} = 0 \ , \tag{8.3}$$

où $\mathcal{U} := U^\Omega$ est l'ensemble des contrôle feedback et $[A(\mathbf{u})v](x) = A(\mathbf{u}(x))v(x)$.

On suppose donnée une discrétisation de $A(u)v$ qui conserve le principe du maximum. Ceci signifie qu'on suppose donnés, pour tout pas de discrétisation h, un ensemble fini Ω_h approchant Ω (par exemple $\Omega \cap (h\mathbb{Z})^d$), et pour tout contrôle $u \in U$, et tout $x \in \Omega_h$, une expression $A_h(u)v(x)$ discrétisant $A(u)v(x)$, et une discrétisation $c_h(u)(x)$ du coût instantané $c(u,x)$, tels que si $\mathbf{u} \in \mathcal{U}_h := U^{\Omega_h}$ est un feedback, alors $A_h(\mathbf{u})v \leq 0 \;\Rightarrow\; v \geq 0$, les inégalités étant pour l'ordre partiel produit de $\mathbb{R}^{\Omega_h}$. Cette propriété est appelée *principe du maximum discret*. Par exemple, dans le cas où les matrices $(a_{ij}(x))$ sont diagonales, on peut utiliser une discrétisation aux différences finies classique des dérivées partielles d'ordre 2, et décentrée des dérivées partielles d'ordre 1. Pour le cas général, voir par exemple [KD92] ainsi que quelques développements récents [BZ03, MZ05].

La discrétisation de (8.1) s'écrit alors :

$$\inf_{u\in U} \{[A_h(u)\, v + c_h(u)](x)\} = 0 \quad x \in \Omega_h \ ,$$

ou vectoriellement :

$$\inf_{\mathbf{u}\in\mathcal{U}_h} \{A_h(\mathbf{u})\, v + c_h(\mathbf{u})\} = 0 \ . \tag{8.4}$$

En général cette équation peut aussi s'écrire sous la forme d'une équation de point fixe :

$$v = \inf_{\mathbf{u}\in\mathcal{U}_h} \{M_h(\mathbf{u})\, v + \tilde{c}_h(\mathbf{u})\} \ , \tag{8.5}$$

où $M_h(\mathbf{u})$ est sous-markovienne pour tout feedback $\mathbf{u}$. Cette équation s'interprète donc comme l'équation de la programmation dynamique d'un problème de contrôle stochastique à temps discret et à espace d'état fini.

Les méthodes classiques pour résoudre ce type d'équation sont l'itération sur les valeurs ou de point fixe sur l'équation (8.5), et l'itération sur les politiques ou algorithme de Howard. Pour la discrétisation d'une équation elliptique telle que (8.1a), la vitesse de convergence de l'itération sur les valeurs est du même ordre que celle de l'algorithme de Jacobi pour les discrétisations d'équations elliptiques linéaires, i.e. $1 - O(h^2)$. Ainsi, le temps de calcul total pour arriver à une erreur de l'ordre de l'erreur de discrétisation (h^θ pour un certain $\theta > 0$) est de l'ordre de $-(\log(h)N_h)/h^2$, ce qui est de l'ordre de $N_h^{1+2/d}\log(N_h)$ pour une discrétisation différences finies de pas h, si N_h est le nombre de points de discrétisation.

Rappelons que l'algorithme d'itérations sur les politiques consiste à calculer les contrôles feedbacks $u_h^k \in \mathcal{U}_h$ et les fonctions valeurs $v_h^k \in \mathbb{R}^{\Omega_h}$ par la récurrence (partant de v_h^0) :

$$u_h^{k+1}(x) \in \operatorname*{Argmin}_{u\in U} [A_h(u)\, v_h^k + c_h(u)](x) \quad \forall x \in \Omega_h \tag{8.6a}$$

$$v_h^{k+1} \text{ est solution de } A_h(u_h^{k+1})\, v + c_h(u_h^{k+1}) = 0 \ . \tag{8.6b}$$

Le principe du maximum assure la convergence décroissante, lorsque k tend vers l'infini, de v_h^k vers la solution v_h de (8.4), mais le taux de convergence peut dépendre en général du pas de discrétisation h. De plus le calcul exact de la solution de (8.6b) est coûteux, de l'ordre de $N_h^{3-2/d}$, si N_h est le nombre de points de discrétisation d'une discrétisation par différences finies standard.

Dans [T1, B1, B2], nous avons introduit l'algorithme *Howard-multigrilles* (MGH) qui consiste à remplacer la résolution exacte de (8.6b) par un certain nombre m_k d'itérations multigrilles, de condition initiale v_h^k.

Rappelons que la méthode multigrille consiste à répartir le calcul de la solution d'une équation sur plusieurs niveaux de grilles, en corrigeant le calcul précédent soit par application de relaxations soit par résolution d'une équation agrégée de l'erreur dans une grille plus grossière. Ainsi, alors que les relaxations telles que l'algorithme de Jacobi ont un taux de convergence de l'ordre de $1-O(h^2)$, la méthode multigrille appliquée à une équation uniformément fortement elliptique a un taux de convergence indépendant de h pour un temps de calcul du même ordre que les relaxations, i.e. N_h.

D'autre part si les problèmes de minimisation sous-jacents à (8.1a) ont de bonne propriétés de convexité, alors le feedback optimal dans (8.3) est régulier par rapport à la fonction v, ce qui implique que l'algorithme de Howard est équivalent à celui de Newton, et donc qu'il a une convergence super-linéaire. Ainsi la combinaison de l'algorithme de Newton-Howard avec les méthodes multigrilles fournit un algorithme global avec un taux de convergence indépendant de h, au moins au voisinage de la solution. Cette combinaison peut être faite de deux manières différentes. Soit l'on résout les équations linéaires sous-jacentes à l'algorithme de Newton par une méthode multigrille, c'est ce que nous faisons dans l'algorithme MGH et qui a été étudié aussi par Bank et Rose [BR82]. Soit l'on résout l'équation non-linéaire par une méthode multigrille non-linéaire où l'erreur de l'équation non-linéaire est approximée au moyen d'une équation non-linéaire "agrégée". Dans le cas des équations d'Hamilton-Jacobi-Bellman, cette méthode a été étudiée par Hoppe [Hop86].

Dans les deux cas la convergence rapide n'est assurée qu'au voisinage de la solution. Nous avons donc introduit l'algorithme FMGH qui combine l'algorithme MGH avec un raffinement successif des grilles, comme dans l'algorithme FMG. Dans [T1], nous avons montré, en utilisant le principe du maximum, que l'algorithme FMGH résout l'équation discrète avec la même précision que celle de la discrétisation et avec un temps de calcul de l'ordre du nombre de pas de discrétisation, ce qui est le mieux que l'on puisse obtenir avec une implémentation séquentielle.

Plus précisément, on se donne des grilles imbriquées $\Omega_0 \subset \cdots \subset \Omega_\ell$ de pas de discrétisation $h_0 > \cdots > h_\ell$. On note $\mathcal{V}_j = \mathbb{R}^{\Omega_j}$, $\mathcal{U}_j = \mathcal{U}_{h_j}$, $A_j = A_{h_j}$. Pour chaque équation linéaire de la forme $A_k(\mathbf{u})v + c_k(\mathbf{u}) = 0$ où $\mathbf{u} \in \mathcal{U}_k$, on peut construire un cycle multigrille associé, dont on note $\mathcal{M}_k(\mathbf{u})$ et $M_k(\mathbf{u})$ les opérateurs affines et linéaires. On introduit de plus les opérateurs d'interpolation $\mathcal{I}_k^{k+1} : \mathcal{V}_k \to \mathcal{V}_{k+1}$.

L'*algorithme FMGH* consiste alors à construire à partir de $v_0^0 \in \mathcal{V}_0$, les suites $v_k^n \in \mathcal{V}_k$ et $u_k^n \in \mathcal{U}_k$ pour $0 \leq n \leq \bar{n}$ et $0 \leq k \leq \ell$ par :

$$
\begin{aligned}
&u_k^{n+1} := u_k(v_k^n) := \operatorname*{Argmin}_{\mathbf{u}\in\mathcal{U}_k}(A_k(\mathbf{u})v_k^n + c_k(\mathbf{u})) \ , \\
&v_k^{n+1} = \mathcal{M}_k^{m_{n+1}}(u_k^{n+1})\, v_k^n \ , \\
&v_{k+1}^0 = \mathcal{I}_k^{k+1}\, v_k^{\bar{n}} \ .
\end{aligned}
$$

On suppose donné un espace de Hilbert $\mathcal{V}$ de type $H^1(\Omega)$, et on note $\|.\|_V$ la norme correspondante. On muni aussi $\mathcal{V}_k$ d'une norme $\|.\|_{V_k}$ telle que $\frac{1}{C}\|v_k\|_{V_k} \leq \|p_k v_k\|_V \leq C\|v_k\|_{V_k}$, où $p_k : \mathcal{V}_k \to \mathcal{V}$ est un opérateur d'interpolation.

Théorème 8.1 ([B2, Théorème 2],[T1, Théorème 3.3])**.** *Notons v la solution de* (8.1)*, v_h la solution de* (8.4) *et $v_k = v_{h_k}$, $k = 0, \dots, \ell$. Soit v_k^n la suite de l'algorithme FMGH, et supposons :*

- *$A_h(\mathbf{u})$ vérifie le principe du maximum discret, pour tout $\mathbf{u} \in \mathcal{U}_h$ et tout $h \le h_0$;*
- *soit $H(x,u,p) := g(u,x)\cdot p + c(u,x)$, alors $\bar{u}(p,x) \in \mathrm{Argmin}_{u\in U}(H(x,u,p))$ admet une solution unique, et il existe $0 < \beta \le 1$ tel que*

$$\|\bar{u}(p',x) - \bar{u}(p,x)\| \le C_u \|p' - p\|^\beta \qquad \forall p, p' \in \mathbb{R}^d \ ;$$

- *l'opérateur $A_h(\mathbf{u})$ est coercif : $\langle -A_h(\mathbf{u})v, v\rangle \ge C_1 \|v\|_{V_h}^2$ pour tous $h \le h_0$, $\mathbf{u} \in \mathcal{U}_h$ et $v \in \mathcal{V}_h$;*
- *$\|M_k(\mathbf{u})\|_{V_k} \le \rho < 1$ pour tous $\mathbf{u} \in \mathcal{U}_k$ et $k \ge 0$;*
- *$\|\mathcal{I}_k^{k+1}\|_{V_{k+1}\leftarrow V_k} \le C_i$, pour tout $k \ge 0$;*
- *la discrétisation est d'ordre $\theta > 0$: $\|p_h v_h - v\|_V \le C_\theta h^\theta$, pour tout $h \le h_0$.*

Alors pour tout $\eta > 0$ il existe $\underline{m} \ge 1$ tel que si $m_k \ge \underline{m}$ à partir d'un certain rang, alors pour tout $v_0^0 \in \mathcal{V}_0$, il existe $\underline{n}$ tel que si $\bar{n} \ge \underline{n}$ alors

$$\|v_k^{\bar{n}} - v_k\|_V \le \eta C_\theta h_k^\theta \ ,$$
$$\|p_k v_k^{\bar{n}} - v\|_V \le (C\eta + 1) C_\theta h_k^\theta \ .$$

Ce résultat est prouvé dans [T1]. Il faut noter que celui-ci utilise une norme de type H^1 car la convergence des méthodes multigrilles ne peut avoir lieu correctement que dans ce cadre. Il demande donc que la solution v de l'équation (8.1) soit prise dans un sens variationnel, et non de viscosité, ce qui est en général moins bien adapté aux problèmes de contrôle.

L'algorithme FMGH présenté ici a été utilisé dans [C4, J2, J3].

8.2 Extension aux problèmes à horizon fini, ou ergodiques

Nous avons étendu les algorithmes MGH et FMGH au cas de problèmes de contrôle stochastique à horizon fini, dans le cadre d'une discrétisation en temps implicite ou de Cranck-Nicholson. Nous n'avons pas établi de preuve de convergence, mais celle-ci doit pouvoir se faire aisement sous des hypothèses de même type que dans le théorème 8.1. Ces algorithmes ont été utilisés dans [C6, C7].

Nous avons aussi adapté les algorithmes MGH et FMGH au cas de problèmes de contrôle stochastique ergodique. Dans ce cas on doit résoudre une équation de la forme :

$$\inf_{u\in U}\{[A(u)\,v + c(u)](x)\} + \rho = 0 \quad x \in \Omega \tag{8.7}$$

où

$$A(u)\,v(x) = \sum_{ij} a_{ij}(x)\frac{\partial^2 v}{\partial x_i \partial x_j}(x) + \sum_j g_j(u,x)\frac{\partial v}{\partial x_j}(x) \ . \tag{8.8}$$

L'algorithme proposé consiste à modifier alternativement ρ et v, en calculant la nouvelle valeur de v par l'algorithme MGH dans lequel la grille la plus grossière est réduite à un point et la nouvelle valeur de ρ de telle sorte que la moyenne des équations de la version discrétisée de (8.7) soit nulle. Cette méthode ne peut converger que si la solution v de (8.7) est unique (à une constante additive près), mais la convergence reste à établir. L'algorithme a été utilisé dans [J8].

8.3 Développement logiciel

Les applications potentielles principales du contrôle stochastique sont la gestion de stock, la régulation dans les barrages, les finances,... Le système expert *Pandore* développé (de 83 à 90) dans le projet Théosys puis Méta2 de l'INRIA Rocquencourt, sur la machine lisp Symbolics, automatisait la résolution de tout problème de contrôle stochastique (à condition que l'espace des états ne soit pas trop compliqué) : choix des méthodes numériques, génération du programme Fortran de résolution numérique, génération des courbes et du rapport en LaTeX avec références et résolution théorique (existence et unicité) de l'équation. Pandore a été écrit en plusieurs langages qui interagissent entre eux sur la machine Symbolics : Lisp, Macsyma, Prolog.

Les algorithmes numériques MGH et FMGH présentés dans les sections précédentes ont fait l'objet d'un générateur de programmes, écrit en Macsyma, et générant des programmes Fortran. Ce générateur de programmes génère un programme adapté à chaque problème de contrôle stochastique, en fonction des données, en utilisant les facilités du calcul formel. Le programme généré résout le problème particulier au moyen de l'algorithme MGH ou FMGH. Cela a permis notamment la réalisation de tests systématiques dont les résultats sont dans [T1].

Ce programme a été utilisé pour la résolution de problèmes de gestion de portefeuille, dont certains sont présentés dans le chapitre suivant.

Le prix élevé de la machine Symbolics, puis la disparition de celle-ci, ainsi que la non existence d'une machine comparable, a limité la diffusion et maintenant l'utilisation de Pandore, et donc de ce programme générateur.

8.4 Perspectives / Travaux en cours

Les travaux présentés dans ce chapitre sont anciens, et les programmes correspondants aux algorithmes ne peuvent plus être utilisés, pour les raisons indiquées plus haut. Pourtant nous aimerions reprendre ces idées avec des outils plus généraux.

Par exemple les algorithmes MGH et FMGH présentés ici demandent à ce que l'équation discrète à résoudre soit la discrétisation d'une équation aux dérivées partielles. Si on voulait traiter le cas d'une équation de la programmation dynamique directement issue d'un problème de contrôle stochastique à temps discret et espace fini, il faudrait remplacer les opérateurs $A_k(\mathbf{u})$ pour $k = 0, \ldots, \ell-1$ par des "agrégations" de l'opérateur $A_\ell(\mathbf{u})$ obtenus par exemple au cours de la méthode multigrille algébrique [RS87] (voir aussi [VBM01] pour des versions plus récentes de cette méthode) et la méthode multigrille par la méthode multigrille algébrique. Cette idée a été utilisée par exemple en connection avec des méthodes d'apprentissage en contrôle stochastique dans [ZS05].

Une autre généralisation peut être faite en utilisant les itérations sur les politiques adaptées aux problèmes ergodiques de contrôle stochastique ou aux problèmes de jeux à somme nulle [CTG06], ce qui permettrait en particulier de résoudre les problèmes ergodiques lorsque la solution v de (8.7) n'est pas unique (à une constante additive près).

Chapitre 9

Étude de quelques problèmes de contrôle optimal stochastique intervenant en gestion de portefeuille

Les travaux présentés ici ont été obtenus en collaboration avec Agnès Sulem (INRIA Rocquencourt), José Luis Menaldi (Wayne State University), Michael Taksar (University of Missouri), Pierre Séquier (CCF, au moment des travaux), Adnan Aboulalaa (stagiaire d'option de l'École Polytechnique au moment des travaux). Ils ont fait l'objet des articles publiés [J2, J3, J8, C4, C6, C7, C8].

Nous nous intéressons ici à la résolution théorique et numérique de problèmes de gestion de portefeuille avec plusieurs actifs risqués et des coûts de transaction proportionnels. L'équation de la programmation dynamique induite par ces problèmes est une inéquation variationnelle, la frontière libre délimitant la zone de non-transaction. Les contrôles représentant les achats et ventes d'actions sont singuliers. La résolution numérique a été faite au moyen de programmes générés (cf. la section 8.3) comprenant l'algorithme MGH ou FMGH, décrit plus haut. La résolution théorique fait appel aux techniques de solutions de viscosité à la Crandall-Ishii-Lions.

9.1 Le modèle général

Les trois problèmes étudiés dans [J2, J3, J8, C4, C6, C7, C8] suivent à peu près le même modèle. On considère un investisseur ayant à sa disposition un compte en banque à taux fixe et $n \geq 1$ actifs risqués de loi log-normale. Alors si l'on note $s_0(t)$ (resp. $s_i(t)$) la quantité d'argent dans le compte en banque (resp. le i-ième actif risqué), on aura en l'absence de transaction et de consommation :

$$\begin{aligned} ds_0(t) &= rs_0(t)dt \ , \\ ds_i(t) &= \alpha_i s_i(t)dt + s_i(t)dW_i(t) \ , \end{aligned}$$

où W est un Brownien n dimensionnel corrélé et non normalisé, vérifiant : $\mathbb{E}(W_i(t)W_j(t)) = a_{ij}t$.

Si on note c la (densité de) consommation éventuelle, $\mathcal{L}_i(t)$ et $\mathcal{M}_i(t)$ les prix cumulés des achats et des ventes de l'actif i sur l'intervalle de temps $[0,t]$ et λ_i et μ_i les coût proportionnels respectifs lors de l'achat et de la vente de l'actif i, les processus contrôlés s_0 et s_i vérifient alors :

$$
\begin{aligned}
ds_0(t) &= (rs_0(t) - c(t))dt + \sum_{i=1}^{n}(-(1+\lambda_i)d\mathcal{L}_i(t) + (1-\mu_i)d\mathcal{M}_i(t)), \\
ds_i(t) &= \alpha_i s_i(t)dt + s_i(t)dW_i(t) + d\mathcal{L}_i(t) - d\mathcal{M}_i(t), \quad i = 1, \ldots, n.
\end{aligned}
$$

Une politique d'investissement consiste en un processus adapté $(c(t), (\mathcal{L}_i(t), \mathcal{M}_i(t))_{i=1,\ldots,n})$ tel que $\int_0^t c(s,\omega)ds < \infty$ pour presque tout (t,ω), et tel que $\mathcal{L}_i(t)$, $\mathcal{M}_i(t)$ soient continues à droite, croissantes et vérifient $\mathcal{L}_i(0^-) = \mathcal{M}_i(0^-) = 0$.

A chaque instant t le capital net de l'investisseur $\mathcal{W}(t)$ est par définition l'argent qu'il lui resterait sur son compte en banque si il ramenait ses comptes en actifs risqués à 0, par des achats ou ventes d'actions. On a alors $\mathcal{W}(t) = \rho(s(t))$ où

$$
\rho(x) = x_0 + \sum_{i=1}^{n} \min((1-\mu_i)x_i, (1+\lambda_i)x_i)
$$

et la région de solvabilité est :

$$
\mathcal{S} = \{x = (x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1}, \quad \rho(x) \geq 0\} . \tag{9.1}
$$

Par contre si on n'autorise ni crédit ni emprunt sur chacun des comptes, le domaine du problème sera plutôt

$$
\mathcal{S} = \mathbb{R}_+^{n+1} . \tag{9.2}
$$

Dans l'un ou l'autre cas, une politique d'investissement sera dite *admissible* si $s(t)$ reste dans $\mathcal{S}$.

Ce type de modèle a été étudié d'abord par Merton [Mer71] dans le cas sans coûts de transaction. David et Norman [DN90] ont étudié le cas avec coûts de transaction proportionnels, consommation et un actif. Une nombreuse littérature a suivit. Citons en particulier [FS93] où l'on trouve des exemples et références pour l'approche par solutions de viscosité.

9.2 Un modèle avec maximisation de la consommation

Dans [C4, J2, J3], nous avons considéré le problème de la maximisation d'une fonction d'utilité actualisée de la consommation :

$$
V(x) = \sup \mathbb{E}\left[\int_0^\infty e^{-\delta t} u(c(t))dt \mid s(0^-) = x\right] , \tag{9.3}
$$

où $\delta > 0$ est le taux d'actualisation, $u(c)$ est une fonction d'utilité définie par

$$
u(c) = \frac{c^\gamma}{\gamma}, \quad 0 < \gamma < 1 ,
$$

et le supremum dans (9.3) est pris sur toute les politiques d'investissement admissibles, pour $\mathcal{S}$ défini par (9.1).

On obtient alors :

Théorème 9.1 ([J3, Theorem 3.1]). *Supposons que la matrice a des corrélations est diagonale et que*

$$(H1) \qquad \delta > \gamma \left(r + \frac{1}{2(1-\gamma)} \sum_{i=1}^{n} \frac{(\alpha_i - r)^2}{a_{ii}} \right),$$

$$(H2) \qquad 0 \leq \mu_i < 1, \quad \lambda_i \geq 0, \quad \lambda_i + \mu_i > 0, \quad \forall i = 1, \dots, n\ .$$

Alors

(i) la fonction valeur V définie par (9.3) *est Hölder d'ordre γ, concave dans $\mathcal{S}$, et croissante par rapport à chaque variable x_i, $i = 0, \dots n$.*

(ii) V est l'unique solution de viscosité de l'inéquation variationnelle :

$$-\max\left\{ AV + u^\star\Big(\frac{\partial V}{\partial x_0}\Big), \max_{1 \leq i \leq n} L_i V, \max_{1 \leq i \leq n} M_i V \right\} = 0 \quad \text{dans } \mathring{\mathcal{S}} \tag{9.4a}$$

$$V = 0 \quad \text{sur } \partial\mathcal{S} \tag{9.4b}$$

où

$$\begin{aligned} AV &= \frac{1}{2} \sum_{i=1}^{n} a_{ii} x_i^2 \frac{\partial^2 V}{\partial {x_i}^2} + \sum_{i=1}^{n} \alpha_i x_i \frac{\partial V}{\partial x_i} + r x_0 \frac{\partial V}{\partial x_0} - \delta V, \\ L_i V &= -(1 + \lambda_i) \frac{\partial V}{\partial x_0} + \frac{\partial V}{\partial x_i}, \\ M_i V &= (1 - \mu_i) \frac{\partial V}{\partial x_0} - \frac{\partial V}{\partial x_i}, \end{aligned}$$

et où $u^\star$ est la "transformée de Legendre-Fenchel" de u définie par

$$u^\star(\nu) = \max_{c \geq 0}(-c\nu + u(c)) = \Big(\frac{1}{\gamma} - 1\Big)\nu^{\frac{\gamma}{\gamma - 1}}.$$

La zone de $\mathcal{S}$ où le premier terme de (9.4a) est optimal correspond à la zone de non-transaction. Si $L_i V$ (resp. $M_i V$) est optimal, alors il faut acheter (resp. vendre) une quantité d'actif i afin de se retrouver dans la zone de non-transaction.

La fonction V est de plus homogène d'ordre γ : $V(\beta x) = \beta^\gamma V(x)$ pour tout $\beta > 0$ et $x \in \mathcal{S}$, ce qui permet de se ramener à une inéquation variationnelle de dimension n. Pour le traitement numérique on s'est restreint à l'ensemble

$$\mathcal{S}^+ = \{x \in \mathbb{R}^{n+1},\ x_1, \dots, x_n \geq 0,\ x_0 + \sum_{i=1}^{n} (1 - \mu_i) x_i \geq 0\}\ ,$$

ce qui ne change rien à la solution dans les cas intéressants de paramètres α et a où il n'est pas intéressant de ne jamais utiliser un des actifs risqués.

Nous avons obtenu des résultats en dimension 1 et 2 permettant de décrire l'évolution du capital et de la zone de non-transaction en fonction des coûts de transaction.

9.3 Un modèle avec maximisation du capital net

Dans [C6, C7] nous avons considéré le problème à horizon fini et sans consommation de la maximisation du capital net à un instant T donné :

$$V(T,x) = \sup \mathbb{E}\left[\frac{1}{1-\gamma}\mathcal{W}(T)^{1-\gamma} \mid s(0) = x\right] . \tag{9.5}$$

Ici $\gamma \geq 0$ est tel que $\gamma \neq 1$, c'est le facteur d'aversion au risque. On suppose l'ensemble $\mathcal{S}$ des états admissibles donné par (9.2). Nous obtenons un résultat similaire au théorème 9.1, pour une inéquation variationnelle de type parabolique. L'homogénéité de la fonction valeur est encore vérifiée, ici avec un exposant $1-\gamma$, ce qui permet de se ramener à une inéquation variationnelle en dimension n. L'étude numérique s'est concentrée sur les effets des coûts de transaction sur la stratégie optimale.

9.4 Un modèle de maximisation du taux de croissance moyen en temps long du capital

Dans [C8, J8] nous avons considéré le problème sans consommation, de la maximisation du taux moyen en temps long du capital net :

$$\pi = \sup \liminf_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\log \mathcal{W}(T) \mid s(0) = x\right] , \tag{9.6}$$

dans le cas de l'ensemble des états admissibles $\mathcal{S}$ donné par (9.2).

Après un changement de variable faisant descendre en dimension comme dans les sections précédentes (la fonction valeur du problème à horizon fini vérifie l'homogénéité $V(T,\beta x) = \log(\beta) + V(T,x)$), on se ramène à un problème de contrôle stochastique ergodique standard. L'inéquation variationnelle associée est de la forme

$$-\mathcal{F}(D^2\mathcal{V}(y), D\mathcal{V}(y), \pi, y) = 0 \quad \text{dans } \Delta , \tag{9.7}$$

où $\Delta := \{y \in \mathbb{R}_+^n \mid \sum_{i=1}^n y_i \leq 1\}$, et où $\mathcal{F}$ est définie par :

$$\begin{aligned}
\mathcal{F}(D^2V, DV, \pi, y) &= \max\left(BV - \pi + H(y), \mathcal{F}_0(DV, y)\right) \ , \\
BV &= \frac{1}{2}\sum_{i,j=1}^n b_{ij}(y)\frac{\partial^2 V}{\partial y_i \partial y_j} + \sum_{i=1}^n \beta_i(y)\frac{\partial V}{\partial y_i} \ , \\
b(y) &= \sigma(y) a \sigma(y)^T \ , \\
\sigma_{ij}(y) &= y_i(\delta_{ij} - y_j) \ , \\
\beta_i(y) &= y_i \sum_{k=1}^n \left(- \sum_{l=1}^n a_{kl} y_l + \alpha_k - r \right) (\delta_{ki} - y_k) \ , \\
H(y) &= r + \sum_{i=1}^n (\alpha_i - r) y_i - \frac{1}{2}\sum_{i,j=1}^n a_{ij} y_i y_j \ , \\
\mathcal{F}_0(DV, y) &= \sup_{(p,q)} \sum_{i=1}^n \left(p_i(P_i V - \nu_i) + q_i Q_i V\right) \ , \qquad (9.8) \\
P_i V &= \nu_i \sum_{j=1}^n y_j \frac{\partial V}{\partial y_j} + \frac{\partial V}{\partial y_i} \ , \\
Q_i V &= -\frac{\partial V}{\partial y_i} \ , \\
\nu_i &= \frac{\lambda_i + \mu_i}{1 - \mu_i} \ ,
\end{aligned}$$

pour $y = (y_1, \ldots, y_n) \in \Delta$ et où le supremum dans (9.8) est pris sur tous les $p = (p_1, \ldots, p_n), q = (q_1, \ldots, q_n)$ tels que $p_i, q_i \geq 0, \ \sum_{i=1}^n (p_i + q_i) = 1, \ p_i - q_i \geq 0$ si $y_i = 0$ et $\sum_{i=1}^n (q_i - (1 + \nu_i) p_i) \geq 0$ si $\sum_{i=1}^n y_i = 1$.

On obtient ainsi le résultat suivant :

Théorème 9.2 ([J8, Proposition 7.1 et Theorem 7.2])**.** *L'équation* (9.7) *a une solution de viscosité $\mathcal{V}$ concave pour une constante π et pour toute solution de viscosité $\mathcal{V}$ de* (9.7)*, la constante π est égale au supremum dans* (9.6)*.*

Dans le cas de la dimension $n = 1$, on obtient de plus

Théorème 9.3 ([J8, Theorem 8.2])**.** *Si $n = 1$, la solution $\mathcal{V}$ de* (9.7) *est unique à une constante additive près.*

Ce résulat est inspiré des propriétés de structure de l'ensemble des vecteurs propres obtenues dans le cas discret dans la section 6.1. La situation est analogue à celle où le graphe critique n'a qu'une composante connexe, qui correspond en l'occurence à la zone de non transaction.

Il faut noter qu'un problème similaire et dont le précédent est le cas limite pour $\gamma = 1$ est le suivant

$$\pi = \sup \liminf_{T \to \infty} \frac{1}{T(1-\gamma)} \log \mathbb{E}\left[\mathcal{W}(T)^{1-\gamma} \mid s(0) = x\right] \ . \qquad (9.9)$$

Ici $\gamma \geq 0$, $\gamma \neq 1$ est comme dans la section précédente le facteur d'aversion au risque. Lorsque $\gamma \neq 1$, on retrouve une inéquation similaire à (9.7) à laquelle on rajoute un terme quadratique en la dérivée première multiplié par $(1-\gamma)$. Ainsi si $\gamma < 1$, l'inéquation s'interprète encore comme l'équation de la programmation dynamique d'un problème de contrôle stochastique ergodique, mais si $\gamma > 1$, elle s'interprète comme l'équation de la programmation dynamique d'un problème de jeux à somme nulle ergodique. L'étude théorique et numérique en devient dans ce dernier cas plus délicate.

9.5 Perspectives / Travaux en cours

Les modèles d'optimisation de portefeuille tels que ceux présentés ici permettent de donner des exemples d'applications de méthodes nouvelles en contrôle stochastique. Ils ont par exemple été utilisés dans [C18] pour illustrer les résultats de la section 3.4 sur le théorème de Gärtner-Ellis. Ils pourront aussi être utilisés comme illustration de nouvelles méthodes numériques en contrôle stochastique ou jeux à somme nulle ergodiques. Par ailleurs, d'autres modèles de fonction d'utilité, de prix ou de coûts de transaction ont été introduits dans la littérature, fournissant ainsi d'autres exemples d'applications.

A noter aussi qu'une des perspectives de recherche évoquée dans le chapitre 6 est l'obtention de résultats d'unicité du type de celui du théorème 9.3 pour d'autres cas particuliers de problèmes ergodiques.

Quatrième partie

Bibliographie

Liste complète des publications personnelles

Certaines publications peuvent être obtenues en version finale ou préliminaire en pointant sur "pdf" en marge gauche de ces publications. D'autres peuvent être obtenues en version préliminaire sous forme de préprint arXiv ou de rapport de recherche INRIA, en pointant sur le lien correspondant ou en allant sur ma page Web `http ://www-rocq.inria.fr/who/Marianne.Akian/`. Les versions finales ne peuvent être distribuées en général, et ne sont donc pas disponibles sur ma page Web.

Une sélection de publications et pré-publications représentatives des travaux présentés dans ce mémoire a été regroupée dans l'archive compressée (26M) disponible provisoirement dans : `http ://minimal.inria.fr/publis-akian-non-distribuees/publis-akian-hdr.tar.gz`.

Thèse

[T1] M. Akian. *Méthodes multigrilles en contrôle stochastique.* Thèse de Doctorat, Université Paris IX-Dauphine, Paris, 1990. (cf. `http ://www.inria.fr/rrrt/tu-0107.html`).

Articles parus ou acceptés dans des journaux internationaux

[J1] Max Plus[1]. L'algèbre $(\max, +)$ et sa symétrisation ou l'algèbre des équilibres. *C. R. Acad. Sci. Paris Sér. II Méc. Phys. Chim. Sci. Univers Sci. Terre*, 311(4) :443–448, 1990.

[J2] M. Akian, J. L. Menaldi, and A. Sulem. Multi-asset portfolio selection problem with transaction costs. *Math. Comput. Simulation*, 38(1-3) :163–172, 1995. Probabilités numériques (Paris, 1992).

pdf [J3] M. Akian, J. L. Menaldi, and A. Sulem. On an investment-consumption model with transaction costs. *SIAM J. Control Optim.*, 34(1) :329–364, 1996.

pdf [J4] J. P. Quadrat and Max-Plus Working Group[2]. Min-plus linearity and Statistical Mechanics. *Markov Process. Related Fields*, 3(4) :565–587, 1997. Statistical Mechanics of large networks (Rocquencourt, 1996).

[1]Nom collectif donné au groupe de travail en théorie des Systèmes à Evénements Discrets à l'INRIA (Projet Méta2) composé au moment de l'article de M. Akian, G. Cohen, S. Gaubert, R. Nikhoukhah et J.P. Quadrat.

[2]Currently comprising : M. Akian, G. Cohen, S. Gaubert, M. Mc Gettrick, J.P. Quadrat and M. Viot

pdf [J5] M. Akian, R. B. Bapat, and S. Gaubert. Asymptotics of the Perron eigenvalue and eigenvector using max-algebra. *C. R. Acad. Sci. Paris Sér. I Math.*, 327 :927–932, 1998. Voir aussi le Rapport de recherche INRIA 3450.

pdf [J6] M. Akian. Densities of idempotent measures and large deviations. *Trans. Amer. Math. Soc.*, 351(11) :4515–4543, 1999.

pdf [J7] M. Akian and P.-A. Bliman. On super-high frequencies in discontinuous 1st-order delay-differential equations. *J. Differential Equations*, 162(2) :326–358, 2000.

pdf [J8] M. Akian, A. Sulem, and M. Taksar. Dynamic optimization of long-term growth rate for a portfolio with transaction costs and logarithmic utility. *Math. Finance*, 11(2) :153–188, 2001.

pdf [J9] M. Akian and S. Bismuth. Instability of rapidly-oscillating periodic solutions for discontinuous differential delay equations. *Differential Integral Equations*, 15(1) :53–90, 2002.

pdf [J10] M. Akian, P.-A. Bliman, and M. Sorine. Control of delay systems with relay. *IMA J. Math. Control Inform.*, 19(1-2) :133–155, 2002. Special issue on analysis and design of delay and propagation systems.

pdf [J11] M. Akian, S. Gaubert, and V. Kolokoltsov. Invertibility of functional Galois connections. *C. R. Acad. Sci. Paris*, Ser. I 335(11) :883–888, 2002.

pdf [J12] M. Akian and S. Gaubert. Spectral theorem for convex monotone homogeneous maps, and ergodic control. *Nonlinear Analysis. Theory, Methods & Applications*, 52(2) :637–679, 2003.

pdf [J13] M. Akian and I. Singer. Topologies on lattice ordered groups, separation from closed downward sets and conjugations of type Lau. *Optimization*, 52(6) :629–672, 2003.

pdf [J14] M. Akian, R. Bapat, and S. Gaubert. Perturbation of eigenvalues of matrix pencils and the optimal assignment problem. *C. R. Acad. Sci. Paris, Série I*, 339(2) :103–108, 2004. Voir aussi le Rapport de recherche INRIA 5120 et arXiv:math.SP/0402438.

pdf [J15] M. Akian, S. Gaubert, B. Lemmens, and R. Nussbaum. Iteration of order preserving subhomogeneous maps on a cone. *Math. Proc. Cambridge Philos. Soc.*, 140(1) :157–176, 2006. Voir aussi arXiv:math.DS/0410084.

[J16] M. Akian, S. Gaubert, and A. Lakhoua. The max-plus finite element method for solving deterministic optimal control problems : basic properties and convergence analysis. 2006. à paraître dans SIAM J. Control and Opt., voir arXiv:math.OC/0603619.

Articles publiés dans des livres ou chapitres de livres

[B1] M. Akian. Résolution numérique d'équations d'Hamilton-Jacobi-Bellman au moyen d'algorithmes multigrilles et d'itérations sur les politiques. In *Analysis and optimization of systems (Antibes, 1988)*, volume 111 of *Lecture Notes in Control and Inform. Sci.*, pages 629–640. Springer, Berlin, 1988.

pdf [B2] M. Akian. Analyse de l'algorithme multigrille FMGH de résolution d'équations d'Hamilton-Jacobi-Bellman. In *Analysis and optimization of systems (Antibes, 1990)*, volume 144 of *Lecture Notes in Control and Inform. Sci.*, pages 113–122. Springer, Berlin, 1990.

pdf [B3] M. Akian, J.P. Quadrat, and M. Viot. Bellman processes. In *11th International Conference on Analysis and Optimization of Systems : Discrete Event Systems (Sophia-Antipolis, 1994)*, volume 199 of *Lecture Notes in Control and Inform. Sci.*, pages 302–311. Springer, 1994.

[B4] S. Gaubert and Max Plus[3]. Methods and applications of (max, +) linear algebra. In *STACS 97 (Lübeck)*, volume 1200 of *Lecture Notes in Comput. Sci.*, pages 261–282. Springer, Berlin, 1997. Voir aussi le Rapport de recherche INRIA 3088.

pdf [B5] M. Akian, J.-P. Quadrat, and M. Viot. Duality between probability and optimization. In J. Gunawardena, editor, *Idempotency*, Publications of the Isaac Newton Institute, pages 331–353. Cambridge University Press, 1998.

[B6] M. Akian, S. Gaubert, and V. N. Kolokoltsov. Set coverings and invertibility of functional galois connections. In G. L. Litvinov and V. P. Maslov, editors, *Idempotent Mathematics and Mathematical Physics*, Contemporary Mathematics, pages 19–51. American Mathematical Society, 2005. Aussi ESI Preprint 1447 et arXiv:math.FA/0403441.

[B7] M. Akian, S. Gaubert, and C. Walsh. Discrete max-plus spectral theory. In G. L. Litvinov and V. P. Maslov, editors, *Idempotent Mathematics and Mathematical Physics*, Contemporary Mathematics, pages 53–77. American Mathematical Society, 2005. Aussi ESI Preprint 1485 et arXiv:math.SP/0405225.

pdf [B8] M. Akian, R. Bapat, and S. Gaubert. Max-plus algebras. In L. Hogben, editor, *Handbook of Linear Algebra (Discrete Mathematics and Its Applications)*, volume 39. Chapman & Hall/CRC, 2006. Chapter 25.

pdf [B9] M. Akian, S. Gaubert, and L. Ninove. The T-PageRank : a model of self-validating effects of web surfing. In *Positive systems*, volume 341 of *Lecture Notes in Control and Inform. Sci.*, pages 239–246. Springer, Berlin, 2006.

Actes de conférences internationales (avec comité de lecture ou à titre invité)

pdf [C1] M. Akian and A. Bensoussan. On the stochastic Ramsey problem. In *IEEE CDC, Athènes, Grèce*, 1986.

pdf [C2] M. Akian, J.P. Chancelier, and J.P. Quadrat. Dynamic programming complexity and application. In *IEEE CDC, Austin Texas*, 1988.

pdf [C3] Max Plus[1]. Linear systems in (max, +) algebra. In *IEEE CDC, Honolulu, Hawaii*, 1990.

[C4] M. Akian and A. Sulem. Application of "Pandore", an expert system in stochastic control, to portfolio selection problems. In *Artificial Intelligence, Expert Systems and Symbolic Computing (IMACS)*, pages 389–398, Amsterdam, 1992. North Holand.

pdf [C5] Max-Plus Working Group[3], presented by J.-P. Quadrat. Max-plus algebra and applications to system theory and optimal control. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, pages 1511–1522, Basel, 1995. Birkhäuser.

[C6] M. Akian, P. Séquier, A Sulem, and A. Aboulalaa. A finite horizon portfolio selection problem with multi-risky assets and transaction costs : the dynamic asset allocation example. In *Actes Association Française de Finance (Bordeaux, Juin 1995)*, 1995.

pdf [C7] M. Akian, P. Séquier, and A Sulem. A finite horizon multidimensional portfolio selection problem with singular transactions. In *IEEE CDC, New Orléans*, volume 3, pages 2193–2198, 1995.

[3] Currently comprising : M. Akian, G. Cohen, S. Gaubert, J.P. Quadrat and M. Viot

[C8] M. Akian, A Sulem, and M. Taksar. Maximization of a long term growth rate for a mixed portfolio with transaction costs. In *Proceedings of the 17th IFIP Conference*, pages 132–135, July 1995.

[C9] M. Akian, J.-P. Quadrat, and M. Viot. Bellman processes with independent increments. In *Proceedings of the 17th IFIP Conference*, pages 207–210, July 1995.

[C10] M. Akian and P.-A. Bliman. On super-high-frequencies in discontinuous 1st-order delay-differential equations. In *6th IEEE Mediterranean Conference on Control and Systems, Alghero, Italy*, June 1998.

[C11] M. Akian, P.-A. Bliman, and M. Sorine. P.I. control of nonlinear oscillations for a system with delay. In *8th IFAC LSS'98, Patras, Greece*, July 1998.

pdf [C12] K. Blin, M. Akian, F. Bonnans, E. Hoffman, C. Martini, and K. Zeghal. A stochastic conflict detection model revisited. In *Proceedings of AIAA GNC, Denver*, August 2000.

pdf [C13] K. Blin, M. Akian, F. Bonnans, E. Hoffman, and K. Zeghal. A stochastic conflict detection method integrating planned heading and velocity changes. In *IEEE CDC, Sydney*, 2000.

[C14] M. Akian and S. Gaubert. A spectral theorem for convex monotone homogeneous maps. In *Proceedings of the Satellite Workshop on Max-Plus Algebras, IFAC SSSC'01*, Praha, 2001. Elsevier.

[C15] M. Akian, R. Bapat, and S. Gaubert. Generic asymptotics of eigenvalues using min-plus algebra. In *Proceedings of the Satellite Workshop on Max-Plus Algebras, IFAC SSSC'01*, Praha, 2001. Elsevier.

[C16] M. Akian, S. Gaubert, and A. Lakhoua. A max-plus finite element method for solving finite horizon deterministic optimal control problems. In *Proceedings of MTNS'04*, Louvain, Belgique, 2004. Voir aussi le Rapport de recherche INRIA 5163 et arXiv:math.OC/0404184.

pdf [C17] M. Akian, S. Gaubert, and A. Lakhoua. The max-plus finite element method for optimal control problems : further approximation results. In *Proceedings of the joint 44th IEEE Conference on Decision and Control and European Control Conference ECC 2005 (CDC-ECC'05)*, Seville, Espagne, 2005. Voir aussi arXiv:math.OC/0509250.

pdf [C18] M. Akian, S. Gaubert, and V. N. Kolokoltsov. Solutions of max-plus linear equations and large deviations. In *Proceedings of the joint 44th IEEE Conference on Decision and Control and European Control Conference ECC 2005 (CDC-ECC'05)*, Seville, Espagne, 2005. Voir aussi arXiv:math.PR/0509279.

pdf [C19] M. Akian, S. Gaubert, and C. Walsh. How to find horizon-independent optimal strategies leading off to infinity : a max-plus approach. In *Proc. of the 45th IEEE Conference on Decision and Control (CDC'06)*, San Diego, 2006. Voir aussi arXiv:math.OC/0609243.

[C20] M. Akian. Representation of stationnary solutions of Hamilton-Jacobi-Bellman equations : a max-plus point of view. In G .L. Litvinov, V. P. Maslov, and S. N. Sergeev, editors, *Proc. of the International Workshop "Idempotent and Tropical Mathematics and problems of Mathematical Physics", Moscow, Aug. 25-30, 2007*, volume 1, pages 6–11.

Articles soumis

[S1] M. Akian, R. Bapat, and S. Gaubert. Min-plus methods in eigenvalue perturbation theory and generalised Lidskiĭ-Višik-Ljusternik theorem. Rapport de recherche INRIA 5104, Fév 2004. Et arXiv:math.SP/0402090, soumis.

[S2] M. Akian, S. Gaubert, and C. Walsh. The max-plus Martin boundary. Rapport de recherche INRIA 5429, Déc. 2004. Et arXiv:math.MG/0412408, soumis.

[S3] M. Akian, B. David, and S. Gaubert. Un théorème de représentation des solutions de viscosité d'une équation d'Hamilton-Jacobi-Bellman ergodique dégénérée sur le tore, 2007. Soumis.

[S4] M. Akian, S. Gaubert, and L. Ninove. Multiple equilibria of nonhomogeneous Markov chains and self-validating Web rankings, 2007. Soumis.

Articles en préparation

[P1] M. Akian, S. Gaubert, and V. Kolokoltsov. Invertibility of Moreau conjugacies and large deviations. En préparation.

[P2] M. Akian, S. Gaubert, and R. Nussbaum. The Collatz-Wielandt theorem for order-preserving homogeneous maps on cones. En préparation.

[P3] M. Akian, S. Gaubert, and V. Kolokoltsov. On the assignment problem for a countable state space. En préparation.

[P4] M. Akian, S. Gaubert, and A. Guterman. Rank functions for matrices over max-algebras. En préparation.

[P5] M. Akian, S. Gaubert, and R. Nussbaum. Uniqueness of fixed points of nonexpansive semi-differentiable maps. En préparation.

[P6] M. Akian, S. Gaubert, and B. Lemmens. Stable fixed points in discrete convex monotone dynamical systems. En préparation.

[P7] M. Akian, S. Gaubert, and C. Walsh. How to find horizon-independent optimal strategies leading off to infinity : a max-plus approach. Version complète avec preuves, en préparation.

[P8] M. Akian, B. David, and S. Gaubert. A representation theorem for the viscosity solutions of a degenerate ergodic hamilton-jacobi-bellman equation on the torus. En préparation.

[P9] M. Akian, S. Gaubert, and A. Lakhoua. Solving deterministic optimal control problems with the max-plus finite element method : quadratic error estimates. En préparation.

Rapports de recherche n'ayant pas fait l'objet d'une autre publication

[R1] M. Akian. Theory of cost measures : convergence of decision variables. Rapport de recherche INRIA 2611, 1995.

[R2] M. Akian. On the continuity of the Cramer transform. Rapport de recherche INRIA 2841, 1996.

Références citées dans le texte hormis les publications personnelles

[BCJ00] E. N. Barron, P. Cardaliaguet, and R. R. Jensen. Radon-Nikodym theorem in L^∞. *Appl. Math. Optim.*, 42(2) :103–126, 2000.

[BCOQ92] F. Baccelli, G. Cohen, G. J. Olsder, and J.-P. Quadrat. *Synchronisation and Linearity.* Wiley, 1992.

[BDL05] G. Barles and F. Da Lio. On the boundary ergodic problem for fully nonlinear equations in bounded domains with general nonlinear Neumann boundary conditions. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 22(5) :521–541, 2005.

[Bel92] F. Bellalouna. *Un point de vue linéaire sur la programmation dynamique. Détection de ruptures dans le cadre des problèmes de fiabilité.* Thèse, Université Paris-IX Dauphine, Paris, 1992.

[Ben88] A. Bensoussan. *Perturbation methods in optimal control.* Wiley/Gauthier-Villars Series in Modern Applied Mathematics. John Wiley & Sons Ltd., Chichester, 1988. Translated from the French by C. Tomson.

[BJ72] T. S. Blyth and M. F. Janowitz. *Residuation Theory.* Pergamon press, 1972.

[BNS03] A. D. Burbanks, R. D. Nussbaum, and C. T. Sparrow. Extension of order-preserving maps on a cone. *Proc. Roy. Soc. Edinburgh Sect. A*, 133(1) :35–59, 2003.

[BR82] R. E. Bank and D. J. Rose. Analysis of a multilevel iterative method for nonlinear finite element equations. *Math. Comp.*, 39(160) :453–465, 1982.

[But00] P. Butkovič. Simple image set of (max, +) linear mappings. *Discrete Appl. Math.*, 105(1-3) :73–86, 2000.

[BZ03] J. F. Bonnans and H. Zidani. Consistency of generalized finite difference schemes for the stochastic HJB equation. *SIAM J. Numer. Anal.*, 41(3) :1008–1021 (electronic), 2003.

[CDF89] I. Capuzzo-Dolcetta and M. Falcone. Discrete dynamic programming and viscosity solutions of the Bellman equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 6(suppl.) :161–183, 1989. Analyse non linéaire (Perpignan, 1987).

[CDQV83] G. Cohen, D. Dubois, J.-P. Quadrat, and M. Viot. Analyse du comportement périodique des systèmes de production par la théorie des dioïdes. Rapport de recherche 191, INRIA, Le Chesnay, France, 1983. .

[CG79] R. A. Cuninghame-Green. *Minimax Algebra.* Number 166 in Lecture notes in Economics and Mathematical Systems. Springer, 1979.

[CG86] W. Chou and R. B. Griffiths. Ground states of one dimensional systems using effective potentials. *Phys. Rev. B*, 34 :6219–34, 1986.

[CGQ04] G. Cohen, S. Gaubert, and J.-P. Quadrat. Duality and separation theorems in idempotent semimodules. *Linear Algebra and Appl.*, 379 :395–422, 2004. Also arXiv:math.FA/0212294.

[CGQS05] G. Cohen, S. Gaubert, J.P. Quadrat, and I. Singer. Max-plus convex sets and functions. In G. L. Litvinov and V. P. Maslov, editors, *Idempotent Mathematics and Mathematical Physics*, Contemporary Mathematics, pages 105–129. American Mathematical Society, 2005. Also arXiv:math.FA/0308166.

[Con01] G. Contreras. Action potential and weak KAM solutions. *Calc. Var. Partial Differential Equations*, 13(4) :427–458, 2001.

[CTG06] J. Cochet-Terrasson and S. Gaubert. A policy iteration algorithm for zero-sum stochastic games with mean payoff. *C. R. Math. Acad. Sci. Paris*, 343(5) :377–382, 2006.

[DJLC53] M. L. Dubreil-Jacotin, L. Lesieur, and R. Croisot. *Leçons sur la Théorie des Treillis, des Structures Algébriques Ordonnées, et des Treillis géométriques*, volume XXI of *Cahiers Scientifiques.* Gauthier Villars, Paris, 1953.

[DM94] P. Del Moral. *Résolution particulaire des problèmes d'estimation et d'optimisation non-linéaires.* Thèse, Université Paul Sabatier, Toulouse, 1994.

[DM98] Pierre Del Moral. Maslov optimization theory : topological aspects. In *Idempotency (Bristol, 1994)*, volume 11 of *Publ. Newton Inst.*, pages 354–382. Cambridge Univ. Press, Cambridge, 1998.

[DN90] M. H. A. Davis and A. R. Norman. Portfolio selection with transaction costs. *Math. Oper. Res.*, 15(4) :676–713, 1990.

[DSDM98] B. De Schutter and B. De Moor. The QR decomposition and the singular value decomposition in the symmetrized max-plus algebra. *SIAM J. Matrix Anal. Appl.*, 19(2) :378–406 (electronic), 1998.

[DZ93] A. Dembo and O. Zeitouni. *Large deviations techniques and applications.* Jones and Barlett, Boston, MA, 1993.

[Fat97] A. Fathi. Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(9) :1043–1046, 1997.

[Fat03] A. Fathi. Weak KAM theorem in Lagrangian dynamics. Lecture notes, fourth preliminary version (Cambridge University Press, to appear.), October 2003.

[Fle04] W. H. Fleming. Max-plus stochastic processes. *Appl. Math. Optim.*, 49(2) :159–181, 2004.

[FM00] W. H. Fleming and W.M. McEneaney. A max-plus based algorithm for an hjb equation of non-linear filtering. *SIAM J. Control and Opt.*, pages 683–710, 2000.

[FS93] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Applications of Mathematics (New York).* Springer-Verlag, New York, 1993.

[FS04] A. Fathi and A. Siconolfi. Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation. *Invent. Math.*, 155(2) :363–388, 2004.

[FSF96] L. Fridman, E. Shustin, and E. Fridman. Steady modes and sliding modes in the relay control systems with time delay. In *Proceedings of the 35th Conf. on Decision and Control*, pages 4601–4606, Kobe, Japan, 1996.

[Gau92] S. Gaubert. *Théorie des systèmes linéaires dans les dioïdes.* Thèse, École des Mines de Paris, Juillet 1992.

[GB99] S. Gaubert and P. Butkovic. Sign-nonsingular matrices and matrices with unbalanced determinant in symmetrised semirings. *Linear Algebra Appl.*, 301(1-3) :195–201, 1999.

[GHK+80] G. Gierz, K. H. Hoffman, K. Keimel, J. D. Lawson, M. Mislove, and D. S. Scott. *A compendium of continuous lattices.* Springer Verlag, Berlin, 1980.

[GKZ94] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky. *Discriminants, resultants, and multidimensional determinants.* Birkhäuser, 1994.

[GM77] M. Gondran and M. Minoux. Valeurs propres et vecteurs propres dans les dioïdes et leur interprétation en théorie des graphes. *EDF, Bulletin de la Direction des Etudes et Recherches, Serie C, Mathématiques Informatique*, 2 :25–41, 1977.

[GM78] M. Gondran and M. Minoux. L'indépendance linéaire dans les dioïdes. *E.D.F., Bulletin de la Direction des Etudes et recherches, Série C, Mathématiques, Informatique*, 1 :67–90, 1978.

[GM79] M. Gondran and M. Minoux. *Graphes et algorithmes.* Eyrolles, Paris, 1979. Engl. transl. *Graphs and Algorithms*, Wiley, 1984.

[GM02] M. Gondran and M. Minoux. *Graphes, Dioïdes et semi-anneaux.* TEC & DOC, Paris, 2002.

[Gon96] M. Gondran. Analyse MINPLUS. *C. R. Acad. Sci. Paris Sér. I Math.*, 323(4) :371–375, 1996.

[Hop86] Ronald H. W. Hoppe. Multigrid methods for Hamilton-Jacobi-Bellman equations. *Numer. Math.*, 49(2-3) :239–254, 1986.

[HOvdW05] B. Heidergott, G. J. Olsder, and J. van der Woude. *Max Plus at Work : Modeling and Analysis of Synchronized Systems : A Course on Max-Plus Algebra and It's Applications.* Princeton University Press, 2005.

[KD92] H. J. Kushner and P. G. Dupuis. *Numerical methods for stochastic control problems in continuous time*, volume 24 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1992.

[KM87] V. N. Kolokoltsov and V. P. Maslov. The general form of the endomorphisms in the space of continuous functions with values in a numerical commutative semiring (with the operation $\oplus = \max$). *Dokl. Akad. Nauk SSSR*, 295(2) :283–287, 1987. Engl. transl. in Sov. Math. Dokl., 36 (1), 55-59 (1988).

[KM97] V. N. Kolokoltsov and V. P. Maslov. *Idempotent analysis and applications.* Kluwer Acad. Publisher, 1997.

[KR50] M. G. Kreĭn and M. A. Rutman. Linear operators leaving invariant a cone in a Banach space. *Amer. Math. Soc. Translation*, 1950(26) :128, 1950.

[Lan67] E. Lanery. Étude asymptotique des systèmes markoviens à commande. *Rev. Française Informat. Recherche Opérationnelle*, 1(5) :3–56, 1967.

[LMS01] G. L. Litvinov, V. P. Maslov, and G. B. Shpiz. Idempotent functional analysis : an algebraic approach. *Math. Notes*, 69(5) :696–729, 2001. Also arXiv:math.FA/0009128.

[LS05] B. Lemmens and M. Scheutzow. On the dynamics of sup-norm non-expansive maps. *Ergodic Theory Dynam. Systems*, 25(3) :861–871, 2005.

[Mas73] V. P. Maslov. *Méthodes Operatorielles*. Edition Mir, Moscou, 1973. trad. fr. 1987.

[MBO97] J. Moro, J. V. Burke, and M. L. Overton. On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure. *SIAM J. Matrix Anal. Appl.*, 18(4) :793–817, 1997.

[ME98] Y. Ma and A. Edelman. Nongeneric eigenvalue perturbations of Jordan blocks. *Linear Algebra Appl.*, 273 :45–63, 1998.

[Mer71] R. C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *J. Econom. Theory*, 3(4) :373–413, 1971.

[Mik04] G. Mikhalkin. Amoebas of algebraic varieties and tropical geometry. In *Different faces of geometry*, volume 3 of *Int. Math. Ser. (N. Y.)*, pages 257–300. Kluwer/Plenum, New York, 2004. Also arXiv:math.AG/0403015.

[MLRS02] J.-E. Martínez-Legaz, A. M. Rubinov, and I. Singer. Downward sets and their separation and approximation properties. *J. Global Optim.*, 23(2) :111–137, 2002.

[MLS90] J.-E. Martínez-Legaz and I. Singer. Dualities between complete lattices. *Optimization*, 21(4) :481–508, 1990.

[MPN02] J. Mallet-Paret and R. D. Nussbaum. Eigenvalues for a class of homogeneous cone maps arising from max-plus operators. *Discrete Contin. Dyn. Syst.*, 8(3) :519–562, 2002.

[MPN03] J. Mallet-Paret and R. D. Nussbaum. Boundary layer phenomena for differential-delay equations with state-dependent time lags. III. *J. Differential Equations*, 189(2) :640–692, 2003.

[MS92] V. P. Maslov and S. N. Samborskiĭ. *Idempotent analysis*, volume 13 of *Advances In Soviet Mathematics*. Amer. Math. Soc., Providence, 1992.

[Mur90] K. Murota. Computing Puiseux-series solutions to determinantal equations via combinatorial relaxation. *SIAM J. Comput.*, 19(6) :1132–1161, 1990.

[MZ05] R. Munos and H. Zidani. Consistency of a simple multidimensional scheme for Hamilton-Jacobi-Bellman equations. *C. R. Math. Acad. Sci. Paris*, 340(7) :499–502, 2005.

[Naj99] B. Najman. The asymptotic behavior of the eigenvalues of a singularly perturbed linear pencil. *SIAM J. Matrix Anal. Appl.*, 20(2) :420–427 (electronic), 1999.

[NS01] R. D. Nussbaum and E. Shustin. Nonexpansive periodic operators in l_1 with application to superhigh-frequency oscillations in a discontinuous dynamical system with time delay. *J. Dynam. Differential Equations*, 13(2) :381–424, 2001.

[NS03] A. Neyman and S. Sorin, editors. *Stochastic games and applications*, volume 570 of *NATO Science Series C : Mathematical and Physical Sciences*, Dordrecht, 2003. Kluwer Academic Publishers.

[Nus88] R. D. Nussbaum. Hilbert's projective metric and iterated nonlinear maps. *Memoirs of the AMS*, 75(391), 1988.

[Nus90] R. D. Nussbaum. Omega limit sets of nonexpansive maps : finiteness and cardinality estimates. *Differential Integral Equations*, 3(3) :523–540, 1990.

[OV91] G. L. O'Brien and W. Vervaat. Capacities, large deviations and loglog laws. In S. Cambanis, G. Samorodnitsky, and M.S. Taqqu, editors, *Stable processes and related topics*, volume 25 of *Progress in probability*, pages 43–83. Birkhaüser, 1991.

[Puh01] A. Puhalskiĭ. *Large Deviations and Idempotent Probability.* Number 119 in Monographs and Surveys in Pure and Applied Mathematics. Chapman & Hall, 2001.

[Qua90] J.-P. Quadrat. Théorèmes asymptotiques en programmation dynamique. *C. R. Acad. Sci. Paris Sér. I Math.*, 311(11) :745–748, 1990.

[RGJ95] A. M. Rubinov, B. M. Glover, and V. Jeyakumar. A general approach to dual characterizations of solvability of inequality systems with applications. *J. Convex Anal.*, 2(1-2) :309–344, 1995.

[Rom67] I. V. Romanovskiĭ. Optimization of stationary control of discrete deterministic process in dynamic programming. *Kibernetika*, 3(2) :66–78, 1967.

[Rom73] I. V. Romanovsky. On the solvability of Bellman's functional equation for a Markovian decision process. *J. Math. Anal. Appl.*, 42 :485–498, 1973. Collection of articles dedicated to Salomon Bochner.

[RS87] J. W. Ruge and K. Stüben. Algebraic multigrid. In *Multigrid methods*, volume 3 of *Frontiers Appl. Math.*, pages 73–130. SIAM, Philadelphia, PA, 1987.

[RS01] D. Rosenberg and S. Sorin. An operator approach to zero-sum repeated games. *Israel J. Math.*, 121 :221–246, 2001.

[Sam02] S. N. Samborskiĭ. Extensions of differential operators and nonsmooth solutions of differential equations. *Kibernet. Sistem. Anal.*, (3) :163–180, 192, 2002.

[SF77] P. J. Schweitzer and A. Federgruen. The asymptotic behavior of undiscounted value iteration in Markov decision problems. *Math. Oper. Res.*, 2(4) :360–381 (1978), 1977.

[SF78] P. J. Schweitzer and A. Federgruen. The functional equations of undiscounted Markov renewal programming. *Math. Oper. Res.*, 3(4) :308–321, 1978.

[Shu95] E. Shustin. Super-high-frequency oscillations in a discontinuous dynamic system with time delay. *Israel J. Math.*, 90(1-3) :199–219, 1995.

[Sin90] R. Sine. A nonlinear Perron-Frobenius theorem. *Proc. Amer. Math. Soc.*, 109(2) :331–336, 1990.

[Sin97] I. Singer. *Abstract convex analysis.* John Wiley & Sons Inc., New York, 1997.

[Smi95] H. L. Smith. *Monotone dynamical systems*, volume 41 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 1995. An introduction to the theory of competitive and cooperative systems.

[SS04] D. Speyer and B. Sturmfels. The tropical Grassmannian. *Adv. Geom.*, 4(3) :389–411, 2004.

[Var84] S. R. S. Varadhan. *Large Deviations and Applications*, volume 46 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* SIAM, Philadelphia, Penn., 1984.

[VBM01] P. Vaněk, M. Brezina, and J. Mandel. Convergence of algebraic multigrid based on smoothed aggregation. *Numer. Math.*, 88(3) :559–579, 2001.

[Vir01] O. Viro. Dequantization of real algebraic geometry on logarithmic paper. In *European Congress of Mathematics, Vol. I (Barcelona, 2000)*, volume 201 of *Progr. Math.*, pages 135–146. Birkhäuser, Basel, 2001. Also arXiv:math.AG/0005163.

[Vor67] N. N. Vorobyev. Extremal algebra of positive matrices. *Elektron. Informationsverarbeit. Kybernetik*, 3 :39–71, 1967. in russian.

[Wal05] C. Walsh. Minimum representing measures in idempotent analysis, 2005. Also arXiv:math.MG/0503716.

[Wal07] C. Walsh. The horofunction boundary of finite-dimensional normed spaces. *Math. Proc. Cambridge Philos. Soc.*, 142(3) :497–507, 2007.

[Zim76] K. Zimmermann. *Extremální Algebra*. Ekonomický ùstav ČSAV, Praha, 1976. (in Czech).

[Zim81] U. Zimmermann. *Linear and Combinatorial Optimization in Ordered Algebraic Structures*. North Holland, 1981.

[ZS05] O. Ziv and N. Shimkin. Multigrid methods for policy evaluation and reinforcement learning. In *Proceedings of the 2005 IEEE International Symposium on Intelligent Control (ISIC05)*, pages 1391–1396, Limassol, Cyprus, 2005.