Nonlinear Perron-Frobenius theory and mean-payoff zero-sum stochastic games

Antoine Hochart

NNT : 2016SACLX079

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PARIS-SACLAY
préparée à
L'ÉCOLE POLYTECHNIQUE

ÉCOLE DOCTORALE N°574
Mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques

par

Antoine Hochart

NONLINEAR PERRON-FROBENIUS THEORY AND MEAN-PAYOFF ZERO-SUM STOCHASTIC GAMES

Thèse présentée et soutenue à Palaiseau, le 14 novembre 2016

Composition du jury :

MARIANNE AKIAN	Inria & École polytechnique	codirectrice de thèse
STÉPHANE GAUBERT	Inria & École polytechnique	directeur de thèse
DANIEL HERNÁNDEZ-HERNÁNDEZ	Centro de Investigación en Matemáticas, A.C.	rapporteur
RIDA LARAKI	CNRS & Université Paris-Dauphine	examinateur
JÉRÔME RENAULT	Université Toulouse-Capitole	rapporteur
SYLVAIN SORIN	Université Pierre-et-Marie-Curie	président du jury
WIESŁAW ZIELONKA	Université Paris-Diderot	examinateur

Remerciements

Mes premiers remerciements s'adressent à mes directeurs de thèse Stéphane Gaubert et Marianne Akian, qui ont guidé mes premiers pas dans la recherche. L'enthousiasme et la grande culture mathématique de Stéphane ont été une source d'inspiration et de motivation constante durant ces trois années, quand la grande rigueur, le souci d'exactitude et de précision de Marianne m'incitaient à explorer toujours davantage les concepts et les idées. Je leur suis très reconnaissant de m'avoir fait confiance, et je m'estime très chanceux d'avoir travaillé à leur côté.

Je remercie chaleureusement Daniel Hernández-Hernández et Jérôme Renault d'avoir accepté d'être les rapporteurs de cette thèse. Jérôme avait déjà manifesté son intérêt pour mon travail lors de mon tout premier exposé en tant que doctorant, et c'est un honneur et une grande satisfaction de susciter encore son intérêt.

C'est également un grand honneur de compter parmi les membres du jury Rida Laraki, Sylvain Sorin et Wiesław Zielonka. Je remercie en particulier vivement Rida et Sylvain pour m'avoir fait découvrir les premiers, pendant mon master à Jussieu, la théorie des jeux. Cette thèse doit beaucoup à la passion qu'ils communiquent et la bienveillance dont ils font preuve. Je tiens ici à remercier aussi Hélène Frankowska qui, depuis mon M2, a toujours manifesté à mon égard la même bienveillance.

Pendant ces trois ans, j'ai (parfois) fréquenté le séminaire parisien de théorie des jeux, et son « petit frère », le séminaire des doctorants. Je remercie ses organisateurs, et salue tous les participants qui font vivre ces rencontres. Dans son ensemble, la communauté française de théorie des jeux est très accueillante et sympathique, et c'est toujours un plaisir de fréquenter ses membres : Joon Kwon, Xiaoxi Li, Miquel Oliu Barton, Vianney Perchet, Thomas Rivera, Tristan Tomala, Xavier Venel, Guillaume Vigeral, Bruno Ziliotto, pour n'en citer que quelques-uns.

Préparer une thèse au CMAP, c'est partager la vie du laboratoire. A cet égard, je tiens à remercier ses assistantes et assistants, et particulièrement Nasséra Naar et Alexandra Noiret.

C'est également partager la vie d'une équipe de recherche, en l'occurrence l'équipe Inria Tropical (anciennement Maxplus). J'exprime ma gratitude à l'égard de Jessica Gameiro, assistante de l'équipe, qui s'est toujours montrée agréable et disponible. Je salue aussi ses membres permanents, dont Xavier Allamigeon et Cormac Walsh.

C'est aussi partager la vie des nombreux autres doctorants du CMAP, que je salue (qu'ils me pardonnent de ne pas les nommer). Je remercie en particulier mes compagnons de bureau, présents et passés : Étienne, pour « sa gentillesse, son expertise dans de nombreux domaines, son humour qui fait mouche, et son humilité », mais aussi Vianney, pour tout ça et même plus encore, sans oublier Alexey, Florine, Gustaw, Jean-Bernard, Pascal, Pedro, Simona et Thi Phong.

Enfin, préparer une thèse au CMAP, c'est pouvoir profiter des nombreuses activités proposées à l'École polytechnique, et de l'expertise des personnes qui les animent. Ces trois années auraient été très différentes sans le club Karaté. Je remercie vivement les *sensei* Philippe, Boris et Laurent, mais aussi Ahmed, Hélène et Vincent, entre autre. Un grand merci également à Fabien, dont l'enseignement du Jiu-Jitsu Brésilien n'aura été que trop bref.

Merci à la Fondation Mathématique Jacques Hadamard (FMJH) qui a financé ma thèse.

Pour finir, je remercie affectueusement mes amis, Claire, Jonathan et Sophie, Stefan et Corina ainsi que ma famille.

Résumé (en Français)

Contexte et motivations

Les jeux stochastiques à somme nulle forment une classe de jeux répétés à deux joueurs, introduite par Shapley [Sha53]. Ils décrivent des interactions qui s'étalent sur une période de temps donnée (éventuellement infinie), entre deux agents (appelés « joueurs ») dont les intérêts s'opposent. Ces joueurs prennent des décisions à intervalles de temps fixés, dans le but d'optimiser leur gain respectif sur la durée totale de leur échange. Ce type de jeux apparaît dans plusieurs domaines, dont les sciences informatiques (analyse statique, vérification de modèle), l'économie mathématique, la finance mathématique, ou la dynamique des populations. Le principal trait caractéristique de cette classe de jeux réside dans le fait que les décisions, prises à chaque étape, modifient l'état de la nature selon un processus stochastique contrôlé. Ces décisions génèrent également un paiement instantané d'un joueur vers l'autre, qui dépend en outre de l'état courant de la nature. Ainsi, à chaque étape, les joueurs doivent trouver un compromis entre leur gain instantané et l'évolution de l'état, qui influence les paiements futurs.

Étant donné une évaluation de la suite des paiements générés par les décisions prises par les joueurs à chaque étape, on dit que le jeu à une valeur v si le joueur qui reçoit les paiements (joueur « maximisant ») peut faire en sorte de recevoir au minimum cette quantité v, et si le joueur qui donne les paiements (joueur « minimisant ») peut s'assurer de ne pas perdre plus que cette quantité v, indépendamment de la stratégie suivie par le joueur adverse. Un problème standard est de savoir si la valeur calculée à partir de la moyenne arithmétique des k premiers paiements a une limite quand l'horizon k tend vers l'infini. On appelle cette limite la *valeur limite* (ou encore, la valeur asymptotique). Ce problème a été largement étudié pour différentes classes de jeux répétés, par un grand nombre d'auteurs dont Bewley, Kohlberg, Mertens, Neyman, Sorin, Zamir. Nous renvoyons le lecteur aux articles [MZ72, Koh74, BK76, KN81, MN81, RS01b, Ney03, Sor03]. En particulier, l'existence de la valeur limite des jeux stochastiques dont l'ensemble des états et l'ensemble des actions sont finis a été établie dans [BK76]. Cependant, quand les espaces d'action sont infinis, cette valeur limite peut ne pas exister, même si les données du problèmes sont standards (compacité des espaces d'action, continuité des fonctions de transition) [Vig13].

D'autres problèmes classiques proviennent des *jeux stochastiques à paiement moyen*, pour lesquels la valeur est déterminée à partir de la limite ($\limsup$ ou $\liminf$) des moyennes de Cesàro des paiements d'étapes. Dans ce cas, au lieu d'optimiser sur une période de temps finie, puis de considérer le comportement de la valeur quand les joueurs deviennent de plus en plus patient (horizon $k \to \infty$), on étudie un jeu qui comporte une infinité d'étapes. A l'origine, les jeux à paiement moyen ont été étudiés dans un cadre déterministe par Ehrenfeucht et Mycielski [EM79], et Gurvich, Karzanov er Khachiyan [GKK88]. Les premiers ont montré l'existence de stratégies optimales stationnaires, c'est-à-dire dont les décisions prises par les joueurs à chaque étape dépendent uniquement de l'état courant (et non de l'ensemble des actions passées). Les seconds ont donné un algorithme pour déterminer celles-ci. Cependant, l'existence d'un algorithme en temps polynomial pour résoudre les jeux à paiement moyen demeure, depuis lors, une question ouverte. Il est à noter que ces jeux [ZP96], ainsi que d'autres classes de jeux stochastiques avec états absorbants [Con92] appartiennent à la classe de complexité NP $\cap$ coNP. Ainsi, ils font partie des rares problèmes qui sont dans NP $\cap$ coNP et pour lesquels aucun algorithme en temps polynomial n'est connu. Pour une comparaison des différentes classes de jeux concernées par ces résultats, nous renvoyons à [AM09].

Les jeux stochastiques à paiement moyen soulèvent aussi des problèmes théoriques, tels que l'existence de la valeur ou de stratégies optimales. Ces problèmes ont été étudiés non seulement en théorie des jeux [FV97], mais aussi en théorie du contrôle (c'est-à-dire pour des jeux à un joueur) sous la dénomination de *problèmes de contrôle ergodique*. Ceux-ci sont standards pour les processus de décisions Markoviens [Put94, HLL96], ainsi que pour les problèmes de contrôle stochastique sensible au risque, étudiés en temps discret par plusieurs auteurs dont Hernández-Hernández et Marcus [HHM96], Fleming et Hernández-Hernández [FHH97, FHH99], voir aussi [CCHH05] ou [AB15] pour des résultats plus récents.

Une manière d'analyser les problèmes évoqués ci-dessus est d'exploiter la structure récursive des jeux stochastiques. Celle-ci s'exprime dans leur opérateur de programmation dynamique, appelé *opérateur de Shapley*, dont il est possible de déduire des informations utiles en étudiant leur propriétés analytiques. Cette méthode, appelé « approche opérateur », a été développée par Kohlberg [Koh74, Koh80], Kohlberg et Neyman [KN81], Rosenberg et Sorin [RS01a], Neyman [Ney03], Sorin [Sor04], Renault [Ren12], entre autre. En particulier, la valeur limite des jeux stochastiques ainsi que la valeur des jeux stochastiques à paiement moyen existent et sont indépendantes de l'état initial, si une certaine équation non linéaire aux valeurs propres admet une solution. Cette équation est appelée *équation ergodique* (en Anglais, « *average reward optimality equation* »). Quand l'espace d'état est fini, disons $\{1,\dots,n\}$, cette équation s'écrit

$$T(u) = \lambda e + u \ , \quad \lambda \in \mathbb{R} \ , \ u \in \mathbb{R}^n \ , \tag{0.1}$$

où $T : \mathbb{R}^n \to \mathbb{R}^n$ est l'opérateur de Shapley et e est le vecteur unité de $\mathbb{R}^n$. Le scalaire λ, appelé *constante ergodique*, fournit la valeur pour n'importe quel état initial. En outre, il est possible de déduire du vecteur u, appelé *vecteur de biais*, des stratégies optimales stationnaires. Pour cela, il suffit d'identifier des actions, dans la représentation minimax de T, qui réalisent le maximum et le minimum dans (0.1). Ainsi, comprendre la structure de l'espace des vecteurs de biais est un problème théorique d'une importance fondamentale, puisque cette espace permet d'avoir accès à l'ensemble des stratégies optimales stationnaires.

Dans le cas déterministe à un joueur, la représentation des vecteurs de biais est bien comprise. L'analyse repose sur la théorie spectrale max-plus, dont l'origine remonte aux travaux de Romanovsky [Rom67], Gondran et Minoux [GM77] et Cunnighame-Green [CG79]. Nous renvoyons aussi le lecteur à [MS92, BCOQ92] pour un aperçu plus ample de cette théorie. Quand l'espace d'état et le temps sont continus, une équation aux dérivées partielles de type Hamilton-Jacobi remplace l'équation de programmation dynamique (0.1). Les solutions de cette équation aux dérivées partielles ont été étudiées dans le cadre de la théorie KAM faible, en lien avec des problèmes de contrôle optimal en temps continu, c.f. la monographie de Fathi [Fat08]. Dans le cas stochastique, la structure de l'espace des vecteurs de biais est également bien comprise dans le cadre des problèmes à un joueur en temps discret, c.f. Akian et Gaubert [AG03]. Mais dans le cadre des problèmes à deux joueurs, la structure de cet espace est bien moins comprise, et l'unicité du vecteur de biais, à l'addition d'un vecteur constant près, est déjà un problème important, notamment pour des raisons algorithmiques. En effet, des phénomène de cyclage peuvent se manifester dans l'algorithme d'itération sur les politiques de Hoffman et Karp [HK66], quand l'unicité n'a pas lieu (voir l'exemple dans [ACTDG12]).

Pour résoudre l'équation (0.1), il est habituel de supposer valides des « conditions de communication », basées sur la structure des probabilités de transition. Par exemple, une condition classique impose aux matrices de transition d'être toutes irréductibles, voir [Bat73]. De manière alternative, il est possible de faire appel aux techniques de la théorie de Perron-Frobenius non linéaire. En effet, trouver une solution à l'équation (0.1) est équivalent à trouver un vecteur $u \in \mathbb{R}^n$ dont toutes les coordonnées sont strictement positives, et un scalaire λ strictement positif, tels que l'égalité $f(u) = \lambda u$ soit satisfaite pour une certaine fonction f, définie de l'intérieur du cône standard positif de $\mathbb{R}^n$ dans lui-même, et qui est homogène de degré un et préserve l'ordre partiel standard de $\mathbb{R}^n$.

La théorie de Perron-Frobenius non linéaire s'intéresse plus généralement aux fonctions f, définies d'un cône convexe fermé K d'un espace de Banach dans lui-même, qui sont non expansives par rapport à une métrique induite par K. Selon les propres mots de Nussbaum dans [Nus88], « la question centrale [...] et la difficulté analytique irréductible » est de savoir si f possède un vecteur propre dans l'intérieur de K, c'est-à-dire, si il existe un vecteur $u \in \text{int}(K)$ tel que

$$f(u) = \lambda u \tag{0.2}$$

pour un certain scalaire $\lambda > 0$.

L'intérêt pour l'équation (0.2) remonte aux travaux de Kreĭn et Rutman [KR50]. Elle a été particulièrement étudiée par Nussbaum [Nus88, Nus89], qui a donné des conditions analytiques garantissant l'existence d'une solution. Ces conditions reposent sur la théorie du point fixe. Elle a aussi été étudiée d'un point de vue combinatoire par Gaubert et Gunawardena [GG04], dans le cas où K est l'orthant positif de $\mathbb{R}^n$ et f préserve l'ordre standard associé et est homogène de degré un. Remarquons que ces dernières conditions imposent en particulier que f est non expansive par rapport à la métrique projective de Hilbert. Il est à noter que l'unicité du vecteur propre u, au produit d'un scalaire strictement positif près, est aussi un problème important, qui a été étudié dans les références mentionnées ci-dessus.

Mentionnons que la théorie de Perron-Frobenius non linéaire a des applications dans des domaines aussi variés que la théorie des jeux, la théorie de l'information, la théorie des systèmes dynamiques, la biologie mathématique, l'économie mathématique. En outre, des liens avec les tenseurs positifs ont récemment été mis en lumière (voir le survey [CQZ13]). Nous renvoyons le lecteur à la monographie [LN12] pour un vaste panorama de la théorie de Perron-Frobenius non linéaire.

Dans la présente thèse, nous utilisons des outils de la théorie de Perron-Frobenius non linéaire, introduits dans [Nus88] et [GG04], afin de traiter les problèmes en lien avec l'équation ergodique (0.1), évoqués ci-dessus. Le Chapitre 2 est dédié à la présentation détaillée du cadre de ce travail.

Contributions

Dans cette thèse, nous présentons des résultats nouveaux concernat les jeux stochastiques avec paiement moyen et *espace d'état fini*. L'ensemble du travail exposé dans ce manuscrit repose sur la notion centrale d'« ergodicité » des jeux stochastiques, que nous relions à l'existence d'une solution pour toute une famille de problèmes ergodiques (0.1). Pour caractériser cette notion, nous employons des méthodes de théorie des graphes, impliquant notamment les hypergraphes. Nous étudions aussi l'unicité, à l'addition d'un vecteur constant près, du vecteur de biais. Nous mentionnons que certains des résultats obtenus peuvent se formuler dans le cadre de la théorie de Perron-Frobenius non linéaire, dont les outils sont largement utilisés. Dans les paragraphes suivant, nous passons en revue les principaux résultats.

Dans les Chapitres 4 à 6, nous introduisons et donnons une caractérisation de la notion d'ergodicité pour les jeux stochastiques avec espace d'état fini. Nous empruntons la terminologie au cas des chaînes de Markov.

Selon la définition de Kemeny et Snell [KS76], l'ergodicité d'une chaîne de Markov finie peut être caractérisée par plusieurs propriétés de sa matrice de transition : unicité de la mesure de probabilité invariante ; unicité de la classe finale ; moyennes de Cesàro des paiements obtenus le long d'une trajectoire qui tendent vers une limite indépendante de l'état initial ; existence d'une solution pour toute une famille de problèmes spectraux.

Nous montrons que la plupart de ces caractérisations reste valides dans le cas des jeux stochastiques, avec toutefois une différence majeure concernant la propriété relative à la théorie des graphes : le graphe dirigé associé à la matrice de transition d'une chaîne de Markov finie doit être remplacé par une paire d'hypergraphes dirigés.

Ainsi, nous dirons qu'un jeu, ou son opérateur de Shapley $T : \mathbb{R}^n \to \mathbb{R}^n$, est *ergodique* si l'équation ergodique (0.1) admet une solution pour tous les opérateurs qui s'écrivent sous la forme $g + T$ avec $g \in \mathbb{R}^n$. Ces derniers correspondent à des versions modifiées du jeu original, consistant à ajouter la quantité g_i à tous les paiements dans l'état i. Il s'ensuit que si le jeu est ergodique, la valeur limite est indépendante de l'état initial pour n'importe quelle perturbation additive des paiements.

Dans le Chapitre 4, nous considérons d'abord le cas des jeux dont la fonction de paiement est bornée. Nous montrons, dans le Théorème 4.6, que l'ergodicité est équivalente à l'unicité, à l'addition d'un vecteur constant près, du point fixe d'un opérateur de Shapley auxiliaire. Ce dernier est positivement homogène, c'est-à-dire qu'il commute avec le produit par un scalaire strictement positif. De tels opérateurs apparaissent dans les jeux dont les paiements instantanés sont nuls, et pour cette raison nous les appelons *opérateurs de Shapley sans paiement*.

Cette caractérisation de l'ergodicité nous amène à considérer, dans le Chapitre 5, le problème de l'existence de points fixes d'opérateurs de Shapley sans paiement dont l'$\arg\min$ ou l'$\arg\max$ est fixé. Nous donnons une condition combinatoire en termes de *correspondance de Galois* (Théorème 5.7) ainsi qu'en termes d'*hypergraphes* (Théorème 5.12), et nous proposons un algorithme en temps polynomial (Algorithme 1) qui résout ce problème. Mentionnons que ce problème est intéressant en lui-même puisque la structure de l'ensemble des vecteurs de biais d'un opérateur de Shapley polyédral (associé à un jeu à information parfaite avec espaces d'action finis) est donné localement par l'ensemble des points fixes d'un opérateur de Shapley sans paiement. Nous déduisons ensuite une caractérisation combinatoire de l'ergodicité, qui fait intervenir une paire d'hypergraphes (Corollaire 5.13), dont la construction ne dépend (sous certaines hypothèses) que de la structure du jeu, et non de la valeur que peuvent prendre les différentes données. Traduit dans le langage de la théorie des jeux, nous montrons que l'ergodicité se caractérise par l'absence de « territoires » (dominions, en Anglais) disjoints pour chaque joueur (Théorème 5.16), un territoire étant un ensemble d'état qu'un joueur peut rendre invariant. Nous concluons que vérifier l'ergodicité est coNP-dur, et dépend essentiellement du nombre d'état (Théorème 5.28) : si les états sont fixés, ce problème peut être résolu en temps polynomial par rapport au nombre d'actions.

Nous traitons ensuite le cas général, c'est-à-dire celui des jeux stochastiques dont la fonction de paiement est potentiellement non bornée. D'après un corollaire de Gaubert et Gunawardena [GG04], un jeu stochastique est ergodique, au sens que nous venons de définir, si tous les *espaces de tranche* sont bornés pour la *semi-norme de Hilbert*. Ces espaces sont des ensembles invariants par T, définis pour tout $\alpha, \beta \in \mathbb{R}$ par $\{x \in \mathbb{R}^n \mid \alpha e + x \leqslant T(x) \leqslant \beta e + x\}$. Dans le Chapitre 4, nous montrons que cette condition est non seulement suffisante, mais aussi nécessaire (Théorème 4.21). Pour établir ce résultat, nous employons la théorie des *opérateurs accrétifs*. Plus précisément, nous établissons une condition nécessaire de surjectivité pour de tels opérateurs définis sur un espace normé de dimension finie (Théorème 4.16). En guise de conséquence, nous obtenons un résultat de stabilité, concernant l'existence de points fixes pour toute perturbation additive d'une application non expansive (Corollaire 4.20).

Dans le Chapitre 6, nous caractérisons de manière combinatoire la propriété selon laquelle tous les espaces de tranche sont bornés pour la semi-norme de Hilbert, répondant ainsi à une question ouverte posée dans [GG04]. Cette caractérisation se manifeste par l'absence de territoires disjoints pour chaque joueur dans un jeu stochastique auxiliaire, dont les actions sont définies par le comportement asymptotique de T le long de certaines demi-droites (Théorème 6.2). Nous donnons également une caractérisation en termes d'hypergraphes (Théorème 6.9). Sous l'hypothèse que la fonction de paiement est bornée, cette dernière caractérisation est équivalente à celle fournie dans le Chapitre 5. Quand un seul joueur a des ensembles d'action non triviaux (c'est-à-dire quand T est convexe), nous montrons que ces conditions se simplifient.

Dans les Chapitres 7 et 8 nous étudions l'unicité, à l'addition d'un vecteur constant près, des vecteurs de biais d'un opérateur de Shapley T. Nous commençons par traiter, dans le Chapitre 7, le cas des jeux stochastiques ergodiques à information parfaite et espaces d'action finis, c'est-à-dire quand T est une application polyédrale telle que l'équation ergodique (0.1) admet une solution pour tout opérateur $g + T$ avec $g \in \mathbb{R}^n$. Nous montrons que les vecteurs de perturbation g pour lesquels l'unicité, à l'addition d'un vecteur constant près, n'a pas lieu sont contenus dans les cellules d'un complexe polyédral, dont la codimension est au moins 1 (Théorème 7.8). L'application de ce résultat à l'algorithme d'itération sur les politiques nous permet d'obtenir un schéma de perturbation pour traiter les instances dégénérées de jeux stochastiques (Théorème 7.22).

Ensuite, dans le Chapitre 8, nous généralisons cette propriété « générique » à n'importe quel opérateur de Shapley. Pour cela, nous nous basons sur l'« approche opérateur accrétif » introduite dans le Chapitre 4. Nous montrons que si l'équation ergodique (0.1) admet une solution pour toute perturbation locale de T, alors le vecteur de biais est unique, à l'addition d'un vecteur constant près, pour une perturbation locale « générique » (Théorème 8.6).

Les opérateurs de Shapley définis sur $\mathbb{R}^n$ sont caractérisés par deux propriétés fondamentales : ils préservent l'ordre standard partiel de $\mathbb{R}^n$ et ils sont additivement homogènes, c'est-à-dire qu'ils commutent avec l'addition par un vecteur constant. Cette caractérisation repose sur plusieurs représentations minimax des opérateurs vérifiant ces propriétés, voir [Kol92, RS01b].

Les opérateurs de Shapley qui sont en outre positivement homogènes occupent une place centrale dans la caractérisation de l'ergodicité (Chapitre 4) et sont le principal objet d'étude du Chapitre 5. Dans le Chapitre 3, nous donnons une représentation minimax de tels opérateurs (Corollaire 3.12), justifiant ainsi leur dénomination d'« opérateurs de Shapley sans paiement ».

Pour parvenir à cette représentation, nous montrons qu'une fonction à valeur réelle sur un espace vectoriel topologique est positivement homogène de degré un et non expansive par rapport à une norme faible de Minkowski si, et seulement si, elle peut s'écrire comme le minimax des formes linéaires qui sont non expansives par rapport à cette même norme (Théorème 3.8).

Présentation des chapitres

Chapitre 2

Dans le Chapitre 2, nous introduisons le modèle des jeux stochastiques à somme nulle, nous présentons l'approche opérateur, et nous établissons le lien avec la théorie de Perron-Frobenius non linéaire. En particulier, nous appellerons opérateur de Shapley (abstrait) sur $\mathbb{R}^n$, tout opérateur de $\mathbb{R}^n$ qui est monotone (M) et additivement homogène (AH)

Chapitre 3

Il existe plusieurs théorèmes de représentation minimax des opérateurs de $\mathbb{R}^n$ qui sont monotones (M) et additivement homogènes (AH) : caractérisation en termes de jeux stochastiques [Kol92] (Théorème 2.2), ou en termes de jeux répétés à information parfaite avec transitions déterministes [RS01b, Gun03] (Théorème 2.3). Dans le cadre de la convexité abstraite, mentionnons également [DMLR04, DMLR08], qui fournissent une caractérisation d'opérateurs vérifiant des propriétés similaires dans un cadre multiplicatif. En outre, on obtient le même type de théorèmes de représentation par la dualité de Fenchel-Legendre quand les opérateurs sont en plus convexes (dans ce cas, il n'y a plus d'infimum dans la formule de représentation). En dimension infinie, cela a été appliqué au cas des mesures de risque convexes [FS02] par exemple.

Une sous-classe importante d'opérateurs monotones et additivement homogènes est constituée par ceux qui sont aussi positivement homogènes (PH). C'est notamment le cas des opérateurs de récession, utilisés par Gaubert et Gunawardena [GG04] en théorie de Perron-Frobenius non linéaire. C'est aussi le cas, en dimension infinie, des mesures de risque cohérentes [ADEH99, Del02].

Lorsque $T : \mathbb{R}^n \to \mathbb{R}^n$ est l'opérateur de Shapley d'un jeu stochastique $\Gamma(r)$ dont la fonction de paiement r est bornée, l'opérateur de récession $\widehat{T}$ est bien défini, et est lui-même l'opérateur de Shapley du jeu stochastique $\Gamma(0)$, dont la fonction de paiement est nulle. Cependant, la réciproque ne peut pas être déduite des théorèmes de représentation évoqués ci-dessus : Si $T : \mathbb{R}^n \to \mathbb{R}^n$ est un opérateur de Shapley (abstrait) qui est positivement homogène, ni le Théorème 2.2, ni le Théorème 2.3 ne fournissent une formule de représentation de T comme opérateur de programmation dynamique d'un jeu dont la fonction de paiement est nulle. Dans le Chapitre 3, nous établissons une telle caractérisation. Ce résultat justifie le fait que nous appellerons, dans la suite, *opérateur de Shapley (abstrait) sans paiement* sur $\mathbb{R}^n$, tout opérateur de $\mathbb{R}^n$ qui est monotone (M), additivement homogène (AH) et positivement homogène (PH).

Les deux propriétés qui définissent les opérateurs de Shapley (M) et (AH), sont équivalentes à la propriété d'être non expansif par rapport à une certaine norme faible de Minkowski. Ces dernières sont des substituts de normes qui ne sont pas nécessairement symétriques, ou ne séparent pas nécessairement les points. Elles ont été étudiées en géométrie métrique, en particulier par Papadopoulos et Troyanov [PT14], et apparaissent naturellement dans l'étude des opérateurs de Shapley, voir [GV12].

Pour obtenir une caractérisation minimax des opérateurs monotones, additivement et positivement homogènes, nous commençons par établir un théorème général de représentation minimax qui s'applique à des fonctions réelles définies sur un espace vectoriel topologique, et qui sont non expansives par rapport à une norme faible de Minkowski. Nous caractérisons ensuite les fonctions non expansives qui sont en outre positivement homogènes. Nous en déduisons comme corollaire notre principale application : un théorème de représentation des opérateurs de programmation dynamique des jeux stochastiques dont les paiements de transition sont nuls. Nous en déduisons également qu'il est possible d'approcher ces opérateurs par une application polyédrale positivement homogène. Ce type d'approximation peut être utilisée dans le cadre des méthodes d'éléments finis max-plus pour la résolution numérique d'équations aux dérivées partielles de Hamilton-Jacobi, voir [McE06, AGL08]. Nous déduisons aussi un théorème de représentation de mesures de risque non convexes.

Chapitres 4 à 6

L'existence de la valeur limite d'un jeu stochastique à espace d'état fini est un problème qui peut se résoudre en prouvant l'existence d'une solution à l'équation ergodique (0.1), faisant intervenir l'opérateur de Shapley T du jeu. Pour que cette équation admette une solution, il est nécessaire et suffisant qu'une orbite de T soit bornée pour la semi-norme de Hilbert (Théorème 2.6). Prouver que cette dernière propriété est satisfaite est un problème difficile en général, qui dépend non seulement de la structure du jeu (c'est-à-dire, le support des probabilités de transition), mais aussi de la valeur des différents paramètres (paiements et probabilités de transition). En suivant les idées de Gaubert et Gunawardena [GG04], un problème mieux posé consiste à trouver des conditions pour lesquelles tous les éléments d'une famille donnée d'espaces invariants par T sont bornés pour la semi-norme de Hilbert. De telles conditions ne dépendent habituellement que de la structure du jeu (voir, par exemple, Théorème 2.8). Dans ce cas, l'équation ergodique a une solution non seulement pour T, mais aussi pour toute une famille d'opérateurs perturbés.

Le cas des chaînes de Markov ergodiques est un exemple élémentaire illustrant cette propriété de stabilité. D'après Kemeny et Snell [KS76], une chaîne de Markov finie est ergodique si sa matrice de transition satisfait l'une des assertions listées dans le théorème ci-dessous. Nous rappelons à cet effet que pour une matrice stochastique carrée P de dimension n, le *graphe dirigé* associé est constitué des nœuds $1, \dots, n$ et des arcs (i, j) tels que $P_{ij} > 0$. Une *classe* de la matrice P est un ensemble maximal de nœuds tel que pour n'importe quelle paire de nœuds dans cet ensemble, il existe un chemin dans le graphe de P reliant l'un à l'autre. Une classe est dite *finale* si tout chemin issu d'un nœud de cette classe reste dans celle-ci. Le lecteur pourra trouver la preuve du théorème suivant dans [KS76] et [BP94]

Theorem 0.1. *Soit $P \in \mathbb{R}^{n\times n}$ une matrice stochastique. Les assertions suivantes sont équivalentes :*

(i) tout vecteur $v \in \mathbb{R}^n$ tel que $Pv = v$ est constant ;

(ii) pour tout vecteur $g \in \mathbb{R}^n$, il existe une paire $(\lambda, u) \in \mathbb{R} \times \mathbb{R}^n$ telle que $g + Pu = \lambda e + u$;

(iii) pour tout vecteur $g \in \mathbb{R}^n$, la limite des moyennes de Cesàro $\lim_{k\to\infty}(g + Pg + \cdots + P^{k-1}g)/k$ est un vecteur constant ;

(iv) Le graphe dirigé associé à P a une unique classe finale ;

(v) la matrice stochastique P a une unique mesure de probabilité invariante.

Une chaîne de Markov finie peut être vue comme un jeu stochastique à zéro joueur. Si $P \in \mathbb{R}^{n\times n}$ est sa matrice de transition, et si g_i est le paiement reçu quand l'état $i \in [n]$ est visité, l'opérateur de Shapley s'écrit alors $T(x) = g + Px$ pour tout $x \in \mathbb{R}^n$. On déduit immédiatement que son opérateur de récession est égal à $\widehat{T}(x) = Px$. Ainsi, en utilisant le vocabulaire de la théorie des jeux, l'ergodicité d'une chaîne de Markov finie peut se caractériser soit par le fait que tous les points fixes de l'opérateur de récession $\widehat{T}$ sont constants (Point (i)), soit par le fait que l'équation ergodique a une solution pour tout opérateur $g + T$ avec $g \in \mathbb{R}^n$ (Point (ii)), soit encore par le fait que la valeur limite est constante pour tout vecteur de paiement $g \in \mathbb{R}^n$

(Point (iii)). Une question naturelle est de savoir si ces caractérisations s'étendent aux jeux stochastiques à espace d'état fini.

Rappelons que la stabilité de l'existence d'une valeur limite constante, basée sur des propriétés structurelles, a déjà été étudiée dans plusieurs cadres. Dans le cas des chaînes de Markov finies, pour lesquelles on considère un critère sensible au risque basé sur le coût moyen en temps long, Cavazos-Cadena et Hernández-Hernández [CCHH09] ont donné une condition nécessaire et suffisante sur la matrice de transition pour que l'équation de Poisson – une équation de type programmation dynamique – ait une solution quelle que soit la fonction de coût associée à la chaîne. Toujours motivés par des problèmes dits sensibles au risque, les mêmes auteurs [CCHH10] ont donné, sous une hypothèse de convexité faible, une condition nécessaire et suffisante pour que le problème spectral non linéaire (0.2) ait une solution pour toute perturbation des paiements. En théorie du contrôle optimal (espace d'état et temps continu), Arisawa [Ari97, Ari98] a étudié le problème ergodique pour les équations de Hamilton-Jacobi-Bellman, et a établi un lien entre l'existence de la constante ergodique pour toute fonction de coût ne dépendant que de l'état et l'existence d'un « attracteur ergodique », qui ne dépend que de la dynamique.

Dans les Chapitres 4 à 6, nous étendons la notion d'ergodicité aux jeux stochastiques à somme nulle. Ainsi, nous dirons qu'un jeu (ou de manière équivalente, son opérateur de Shapley $T : \mathbb{R}^n \to \mathbb{R}^n$) est ergodique si l'équation ergodique (0.1) a une solution pour tout opérateur $g + T$ avec $g \in \mathbb{R}^n$.

Dans le Chapitre 4, en nous inspirant des travaux de Gaubert et Gunawardena [GG04], nous commençons par caractériser l'ergodicité en termes d'opérateur de récession. Pour cela, nous supposons que l'opérateur de Shapley et son opérateur de récession ont le même comportement asymptotique. Ensuite, nous relâchons cette hypothèse et caractérisons l'ergodicité en toute généralité, en termes d'espaces de tranche. Ce résultat s'appuie sur le lien entre les applications non expansives et les opérateurs accrétifs.

Dans le Chapitre 5 et le Chapitre 6, nous traitons de l'aspect combinatoire de ces conditions. En particulier, nous les formulons en termes de relation d'accessibilité dans des hypergraphes, ainsi qu'en termes de « territoires » (ou dominions).

Chapitres 7 et 8

La description de l'ensemble des vecteurs de biais d'un jeu stochastique à somme nulle est un problème fondamental. Dans le cas des problèmes à un joueur, c'est-à-dire pour les problèmes de contrôle optimal en temps discret, la représentation des vecteurs de biais et leur relation avec les stratégies optimales sont bien comprises, que les transitions soient déterministes ou stochastiques (processus de décision Markoviens). Quand les transitions sont déterministes, l'analyse de ces problèmes repose sur la théorie spectrale max-plus, qui remonte aux travaux de Romanovsky [Rom67], Gondran et Minoux [GM77], et Cuninghame-Green [CG79]. Pour une présentation de la théorie spectrale max-plus, le lecteur pourra se référer à [MS92, BCOQ92, ABG13]. Kontorer et Yakovenko [KY92] et Kolokoltsov et Maslov [KM97] se sont en particulier intéressés aux problèmes d'optimisation en horizon infini et aux problèmes à paiement moyen. Dans ce cadre, l'ensemble des vecteurs de biais possède la structure d'un cône max-plus, c'est-à-dire qu'il est invariant par combinaisons linéaires max-plus, et il est engendré par une unique famille minimale qui s'identifie au support des mesures maximisantes dans la formulation du problème de contrôle optimal sous forme de programme linéaire. Ces générateurs « extrêmes » correspondent également aux états récurrents des trajectoires optimales infinies. Une interprétation combinatoire de certains de ces résultats, en termes de complexes polyédrales, a récemment été proposée par Sturmfels et Tran [ST13]. L'équation ergodique (0.1) et la structure des vecteurs de biais a aussi été étudié dans le cas d'espaces d'état continus. Un premier cadre d'étude est fourni par la théorie spectrale max-plus en dimension infinie (voir Akian, Gaubert et Walsh [AGW09]), un second par la théorie KAM faible (voir Fathi [Fat08]). En outre, quand les transitions sont stochastiques, la structure de l'ensemble des vecteurs de biais est également bien comprise dans le cas d'espaces d'état finis, voir Akian et Gaubert [AG03].

Dans le cas des jeux à deux joueurs, la structure de l'ensemble des vecteurs de biais est bien moins comprise. Le problème de comprendre quand le vecteur de biais est unique, à une constante additive près, est déjà intéressant, notamment pour des questions algorithmiques. C'est particulièrement le cas pour l'algorithme d'itération sur les politiques de Hoffman et Karp, puisque la non unicité est en général source d'instabilités numériques ou de situations dégénérées. Hoffman et Karp [HK66] ont introduit cet algorithme pour résoudre les jeux stochastiques à somme nulle avec paiement moyen, information parfaite, et espaces d'état et d'action finis. Ils ont montré que l'itération sur les politiques termine si pour chaque paire de stratégie stationnaires des joueurs, la chaine de Markov associée est irréductible. Cependant, des phénomènes de cyclage peuvent apparaître si cette hypothèse d'irréductibilité n'est pas satisfaite, ce qui est le cas pour beaucoup de classes de jeux – en particulier, elle n'est essentiellement jamais satisfaite pour les jeux avec transitions déterministes. Pour contourner cette obstruction, des raffinements ont été proposés par Cochet-Terrasson et Gaubert [CTG06], Akian, Cochet-Terrasson, Detournay et Gaubert [ACTDG12], Bourque et Raghavan [BR14]. Mais le prix à payer pour traiter les situations de non unicité, est une augmentation de la complexité de l'algorithme. Ainsi, il est très intéressant de comprendre quand de tels détails techniques peuvent être évités.

Dans le Chapitre 7, nous considérons le problème de l'unicité du vecteur de biais de jeux stochastiques avec information parfaite et espaces d'état et d'action finis, quand ceux-ci sont ergodiques, c'est-à-dire quand l'équation ergodique (0.1) a une solution pour toute perturbation des paiements de transitions. Le résultat principal (Théorème 7.8) montre que le vecteur de biais est génériquement unique, à l'addition d'un vecteur constant près. Plus précisément, il montre que les vecteurs de perturbation pour lesquels le vecteur de biais n'est pas unique sont contenus dans les cellules d'un complexe polyédrale dont la codimension est au moins 1. Nous déduisons ensuite que l'itération sur les politiques de Hoffman et Karp converge bien pour un paiement générique, ce qui nous amène à proposer un schéma de perturbation explicite qui permet de résoudre les instances non génériques (pour lesquelles des cycles peuvent se produire) sans faire appel à la condition d'irréductibilité.

Dans le Chapitre 8, nous nous occupons du problème de l'unicité du vecteur de biais pour n'importe quel jeu stochastique à somme nulle dont l'espace d'état est fini. Nous montrons que si l'équation ergodique (0.1) a une solution pour toute perturbation locale, alors l'unicité du vecteur de biais est générique localement.

Le Chapitre 3 est basé sur la prépublication [AGH16]. La première partie du Chaptire 4 (concernant les jeux stochastiques avec paiements bornés) et le Chapitre 5 sont issues de l'article [AGH15a], publié dans le journal *Discrete and Continuous Dynamical Systems, Series A*. Les résultats du Chapitre 6 ont été partiellement annoncés dans les actes de la conférence CDC 2015 [AGH15b]. Les résultats du Chapitre 7, exceptés ceux relatifs à l'itération sur les politiques, ont été annoncés dans les actes de la conférence CDC 2014 [AGH14b]. Les résultats se rapportant à l'« approche opérateur accrétif », présentés dans le Chapitre 4 et le Chapitre 8, ont été annoncés dans les actes de la conférence MTNS 2016 [Hoc16].

Contents

CHAPTER 1

Introduction

1.1 Context and motivations

Zero-sum stochastic games are a class of two-player repeated games introduced by Shapley [Sha53]. They describe long-term interactions between two agents (called "players") with opposite interests, which take decisions stage by stage with the objective to optimize their gains over time. They appear in several domains, including computer science, economics or population dynamics. The main feature of this class of games is that the decisions taken at each stage modify the state of nature, according to a controlled stochastic process, and incur an instantaneous payment from one player to the other, which also depends on the current state. Hence, at each stage, the players have to make a trade-off between the instantaneous payment and the evolution of the state which influences future payoffs.

Given an evaluation of the stream of payments, the game has a value v if the maximizing player can be sure to get at least this quantity v, whereas the minimizing player can be sure not to loose more than v, independently of the strategy followed by the other player. A standard problem is to know if the value arising from the arithmetic mean of the first k payments has a limit when the horizon k tends to infinity. We shall call this limit the *mean payoff* (also known as the asymptotic value). This problem has been widely studied for different classes of repeated games, by several authors, including Bewley, Kohlberg, Mertens, Neyman, Sorin, Zamir, see [MZ72, Koh74, BK76, KN81, MN81, RS01b, Ney03, Sor03]. In particular, the existence of the mean payoff of stochastic games with finite state and action spaces was established in [BK76]. However, when the action spaces are infinite, the mean payoff may not exist even with standard assumptions on the data (compactness of the action spaces, continuity of the transition functions) [Vig13].

Other standard problems arise from *mean-payoff stochastic games*, for which the global payoff is the limit ($\limsup$ or $\liminf$) of the Cesàro mean of the payments. Here, instead of optimizing over a finite horizon k and considering the behavior of the value as the players become more patient ($k \to \infty$), we consider a game played in infinitely many stages. Mean-payoff games were originally considered in the deterministic framework by Ehrenfeucht and Mycielski [EM79] and Gurvich, Karzanov and Khachiyan [GKK88]. They respectively showed the existence of optimal stationary strategies, i.e., such that the decisions only depend on the current position, and gave an algorithm to find them. However, the existence of a polynomial-time algorithm to solve mean-payoff games has remained, since then, an open question. Let us note that these games [ZP96], as well as related classes of stopping stochastic games [Con92], were shown to be in the complexity class NP ∩ coNP. Hence, they belong to the few problems that are in NP ∩ coNP for which no polynomial-time algorithm is known. We refer to [AM09] for a comparison of the various classes of games which have been considered. Mean-payoff stochastic games also rise theoretical problems, such as the existence of the value or optimal strategies. These problems have been considered not only in game theory [FV97], but also in control theory (i.e., one-player games) under the name of *ergodic control problems* or *long-run average-reward problems*. The latter problems are standard for Markov decision processes [Put94, HLL96] as well as for risk-sensitive stochastic control, developed in the framework of discrete time by several authors including Hernández-Hernández and Marcus [HHM96], Fleming and Hernández-Hernández [FHH97, FHH99], see also [CCHH05] or [AB15] for more recent developments.

One way to analyze the aforementioned problems is to exploit the recursive structure of stochastic games, encompassed in their dynamic programming operator, so-called *Shapley operator*. One infers useful information by studying the analytical properties of these operators. This method, known as "operator approach", was developed by Kohlberg [Koh74, Koh80], Kohlberg and Neyman [KN81], Rosenberg and Sorin [RS01a], Neyman [Ney03], Sorin [Sor04], Renault [Ren12] and others. In particular, both the mean payoff and the value of mean-payoff games exist and are independent of the initial state if some nonlinear eigenvalue equation, known as *ergodic equation* or *average reward optimality equation*, has a solution. When the state space is finite, say $\{1, \dots, n\}$, this equation writes

$$T(u) = \lambda e + u \ , \quad \lambda \in \mathbb{R} \ , \ u \in \mathbb{R}^n \ , \tag{1.1}$$

where $T : \mathbb{R}^n \to \mathbb{R}^n$ is the Shapley operator and e is the unit vector of $\mathbb{R}^n$. The scalar λ, called *ergodic constant*, gives the value for any initial state, and one can obtain optimal stationary strategies from the vector u, called *bias vector*, by identifying the actions, in the minimax representation of T, that attain the maximum or the minimum in (1.1). Thus, from a theoretical point of view, it is a fundamental issue to understand the structure of the set of bias vectors, since it allows one to access all optimal stationary strategies. In the one-player deterministic case, the representation of bias vectors is well understood. The analysis relies on max-plus spectral theory, which goes back to the work of Romanovsky [Rom67], Gondran and Minoux [GM77] and Cunnighame-Green [CG79], see also [MS92, BCOQ92] for more background on this theory. With continuous time and state space, a Hamilton-Jacobi PDE replaces the dynamic-programming equation (1.1), the solutions of which have been studied in the framework of weak KAM theory, in relation with continuous-time deterministic optimal control problems, see the monograph by Fathi [Fat08]. In the stochastic case, the structure of bias vectors is still well understood for discrete-time one-player problems with finite state space, see Akian and Gaubert [AG03]. But in the two-player case, the structure is less well known and the uniqueness of the bias vector, up to the addition by a constant vector, is already an important issue. The latter problem is also a matter of importance for algorithmic purposes, as nonuniqueness may cause cycling problems in Hoffman-Karp policy iteration algorithm [HK66] (see an example in [ACTDG12]).

The solvability of (1.1) usually requires some "communication conditions" based on the structure of transition probabilities – for instance, all transition matrices being irreducible, see [Bat73]. Alternatively, one can use techniques from nonlinear Perron-Frobenius theory. Indeed, solving Equation (1.1) is the reflection through "$\log$-glasses" of finding a vector $u \in \mathbb{R}^n$

with positive coordinates and a positive scalar λ such that $f(u) = \lambda u$ for some self-map f acting on the standard positive cone of $\mathbb{R}^n$, that is homogeneous of degree one and preserves the standard partial order of $\mathbb{R}^n$.

Nonlinear Perron-Frobenius theory deals more generally with self-maps f acting on a closed convex cone K in a Banach space, that are nonexpansive with respect to some metric induced by K. According to the words of Nussbaum in [Nus88], "the central question [...] and the irreducible analytic difficulty" is to know if f has an eigenvector in the interior of K, i.e., if there exists a vector $u \in \operatorname{int}(K)$ such that

$$f(u) = \lambda u \tag{1.2}$$

for some scalar $\lambda > 0$. The interest for Equation (1.2) goes back to the work of Kreĭn and Rutman [KR50]. It has been especially studied by Nussbaum [Nus88, Nus89], who gave analytic conditions for its solvability, relying on fixed-point theory. It has also been studied from a combinatorial point of view by Gaubert and Gunawardena [GG04] when K is the standard nonnegative cone of $\mathbb{R}^n$ and f is order-preserving and homogeneous of degree one (implying nonexpansiveness with respect to Hilbert's projective metric). Note that the uniqueness of the eigenvector u, up to a positive scalar multiple, is also an important issue, investigated in the latter references. Let us mention that nonlinear Perron-Frobenius theory has applications in various domains, including game theory, information theory, dynamical systems theory, mathematical biology, economics, and links with nonnegative tensors have been recently pointed out (see the survey paper [CQZ13]). We refer to [LN12] for an overview of this theory.

In the present work, we use some tools of nonlinear Perron-Frobenius theory introduced in [Nus88] and [GG04] to address the aforementioned problems related with the ergodic equation (1.1). Chapter 2 is dedicated to the detailed presentation of these topics.

1.2 Contributions

In this thesis, we present new results on mean-payoff stochastic games with finite state space. The whole work is built upon the key notion of *ergodicity* of stochastic games, which we relate with the solvability of a family of ergodic problems (1.1). We make an extensive use of graph-theoretic methods, involving particularly hypergraphs, to characterize the latter notion. We also study the uniqueness, up to the addition by a constant vector, of the bias vector. We mention that some of these results may be transposed to the framework of nonlinear Perron-Frobenius theory, the tools of which are being used to a great extent. We next review the main results.

In Chapters 4 to 6, we introduce and characterize the notion of ergodicity for stochastic games with finite state space. We borrow the term from the Markov chain case. Following the definition of Kemeny and Snell [KS76], the ergodicity of a finite Markov chain may be characterized by several properties involving its transition matrix: uniqueness of the invariant probability measure, uniqueness of the final class, constant limit value of the Cesàro mean of the payments obtained along a trajectory, solvability of a family of spectral problems. We show that most of these characterizations are valid in the case of stochastic games, with however one major discrepancy regarding the graph-theoretic aspect: the directed graph associated with the transition matrix of a finite Markov chain must be replaced by a pair of directed hypergraphs. Thus, we shall say that a game, or its Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$, is *ergodic* if the ergodic equation (1.1) has a solution for all operators $g + T$ with $g \in \mathbb{R}^n$. The latter operators correspond to modified versions of the original game where the quantity g_i is added to all payments in state i. Then, if the game is ergodic, the mean payoff is constant for any such additive perturbation.

In Chapter 4, we first consider the case of games with a bounded payment function. We show in Theorem 4.6 that ergodicity is equivalent to the uniqueness, up to an additive constant, of the fixed point of an auxiliary Shapley operator, which is positively homogeneous, meaning that it commutes with the product by a positive scalar. Such operators arise in games with no instantaneous payments, and for that reason we shall call them *payment-free Shapley operators*. This leads us to consider in Chapter 5 the problem of existence of fixed points with prescribed $\arg\min$ or $\arg\max$ for such operators. We give combinatorial conditions in terms of *Galois connections* (Theorem 5.7) and in terms of *hypergraphs* (Theorem 5.12), and provide a polynomial-time algorithm (Algorithm 1) to solve the latter problem. Note that this problem is interesting in itself since the structure of the set of bias vectors of a polyhedral Shapley operator (game with perfect information and finite action spaces) is given locally by the set of fixed points of some payment-free Shapley operator. We then infer a combinatorial characterization of ergodicity involving a pair of hypergraphs (Corollary 5.13), the construction of which only depends on the structure of the game (under standing assumption). Put in game-theoretic terms, we show that ergodicity is characterized by the absence of disjoint dominions for each player (Theorem 5.16), a dominion being a subset of state controlled by one player. We conclude that checking ergodicity is coNP-hard, but fixed parameter tractable (Theorem 5.28): if the number of states is fixed, the latter problem can be solved in time polynomial with respect to the number of actions.

Then, we address the general case, i.e., games with possibly unbounded payment function. It follows from a corollary of Gaubert and Gunawardena [GG04] that a stochastic game is ergodic, in the sense of our definition, if all *slice spaces* are bounded in *Hilbert's seminorm*. The latter spaces are subsets invariant by T and defined for all $\alpha, \beta \in \mathbb{R}$ by $\{x \in \mathbb{R}^n \mid \alpha e + x \leqslant T(x) \leqslant \beta e + x\}$. In Chapter 4, we show that this condition is not only sufficient but also necessary (Theorem 4.21). To that purpose, we use the theory of *accretive operators*, establishing a necessary condition of surjectivity for any accretive map in a finite-dimensional normed space (Theorem 4.16). As a consequence, we get a stability result (under additive perturbations) for the existence of a fixed point for any nonexpansive map (Corollary 4.20).

In Chapter 6, we provide a combinatorial characterization of the boundedness of all slice spaces in Hilbert's seminorm, answering an open question raised in [GG04]. This characterization is formulated as the nonexistence of disjoint dominions for each player in an auxiliary stochastic game, the action of which are defined by the asymptotic behavior of T along some rays (Theorem 6.2). We also give a graph-theoretic characterization (Theorem 6.9) which recovers the bounded case. When one player is a dummy (i.e., T is convex), we show that these conditions simplify.

In Chapters 7 and 8 we study the uniqueness, up to an additive constant, of the bias vector of a Shapley operator T. In Chapter 7, we first deal with the case of ergodic stochastic games with perfect information and finite action spaces, that is, when T is a polyhedral map such that the ergodic equation (1.1) has a solution for all $g + T$ with $g \in \mathbb{R}^n$. We show that the perturbation vectors g for which uniqueness, up to an additive constant, of the bias vector of $g + T$ fails are contained in the cells of a polyhedral complex, the codimension of which are at least 1 (Theorem 7.8). As an application, we obtain a perturbation scheme allowing one to solve degenerate instances of stochastic games by policy iteration (Theorem 7.22).

Then in Chapter 8, we generalize this "generic property" to any Shapley operator, following the "accretive operator approach" introduced in Chapter 4. We show that if the ergodic equation (1.1) is solvable for local perturbations of T, then the bias vector is unique, up to an additive constant, for a "generic" local perturbation (Theorem 8.6).

Shapley operators on $\mathbb{R}^n$ are characterized by two fundamental properties: they are order-preserving and additively homogeneous, meaning that they commute with the addition by a constant vector. This characterization follows from several minimax representations of operators satisfying the latter properties, see [Kol92, RS01b].

Shapley operators which are also positively homogeneous play a key role in characterizing ergodicity (Chapter 4) and are the central subject of study in Chapter 5. In Chapter 3, we give a minimax representation for such operators (Corollary 3.12), justifying the term of "payment-free Shapley operators". To that purpose we show that a real-valued function on any topological vector space is positively homogeneous of degree one and nonexpansive with respect to a weak Minkowski norm if, and only if, it can be written as a minimax of linear forms that are nonexpansive with respect to the same norm (Theorem 3.8).

Chapter 3 is based on the preprint [AGH16]. The first part of Chapter 4 (regarding stochastic games with bounded payment function) and Chapter 5 are based on the paper [AGH15a], published in *Discrete and Continuous Dynamical Systems, Series A*. The results of Chapter 6 have been partly announced, with a different point of view, in the CDC conference proceedings [AGH15b]. The theoretic aspects of Chapter 7 (not dealing with the application to policy iteration) have been announced in the CDC conference proceedings [AGH14b]. The results of the accretive operator approach, presented in Chapter 4 and Chapter 8, have been announced in the MTNS conference proceedings [Hoc16].

Part I

Operator approach of zero-sum stochastic games

CHAPTER 2

Preliminaries

2.1 Zero-sum stochastic games

In this section, we present the model of zero-sum stochastic, first in its general form (imperfect information), then in the particular case of perfect information. We refer the reader to [FV97, NS03, MSZ14] for more background on these topics.

In the remainder, we shall use the following notation:
- we denote by $\Delta(X)$ the set of probability measures on any measurable space $(X, \mathscr{F})$;
- let $[n] := \{1, \dots, n\}$ for any positive integer n.

2.1.1 General model

Structure of the game

A *(zero-sum) stochastic game* between two players, that we call MIN and MAX respectively, is a class of *zero-sum repeated games* where these two agents, with opposite interests, interact at given time steps. It is defined by a 7-tuple

$$\Gamma := (S, A, B, K_A, K_B, r, p) \ ,$$

where
- S is the *state space*, that we assume finite;
- $(A, \mathscr{A})$ and $(B, \mathscr{B})$ are measurable sets, corresponding to the *action spaces* of players MIN and MAX respectively;
- $K_A \subset S \times A$ and $K_B \subset S \times B$ are measurable sets, representing the *constraint sets* of players MIN and MAX respectively. For each state $s \in S$, the section

$$A_s := \{a \in A \mid (s, a) \in K_A\}$$

is the *set of admissible actions* of player MIN in state s and similarly, the section

$$B_s := \{b \in B \mid (s, b) \in K_B\}$$

is the *set of admissible actions* of player MAX in state s. We assume that all the sets of admissible actions are nonempty, and we let

$$K := \{(s, a, b) \mid s \in S, \ a \in A_s, \ b \in B_s\} \ ,$$

which is a measurable subset of $S \times A \times B$;

– $r : K \to \mathbb{R}$ is a measurable function, representing the *(stage) payment function* (we shall also use the word *payoff function*) for player MAX (and therefore the cost function for MIN);
– $p : K \to \Delta(S)$ is a measurable function, corresponding to the *transition function*.

The game is played in stages, the players having a perfect knowledge of all the past and current information, i.e., they know the previously chosen actions and all the states visited up to the current one. It starts with a given initial state s_1, and then proceeds as follows: at step ℓ, if the current state is s_ℓ, the players choose independently actions a_ℓ and b_ℓ in their respective set of admissible actions, namely A_{s_ℓ} and B_{s_ℓ}. Then, player MAX receives from MIN the (stage) payment $r(s_\ell, a_\ell, b_\ell)$ and the next state $s_{\ell+1}$ is chosen according to the transition probability $p(\cdot \mid s_\ell, a_\ell, b_\ell)$.

In order to avoid technical problems, we shall always assume that for any states $s \in S$ and any probability measures μ and ν over A_s and B_s, respectively, the payoff function $(a, b) \mapsto r(s, a, b)$ is integrable with respect to the product measure of μ and ν. This is in particular the case when r is bounded.

Strategies

Given a finite horizon $k \geqslant 1$, let $\mathscr{H}_k := K^{k-1} \times S$ be the set of admissible *histories* of length k, and let $\mathscr{H}_\infty := K^\infty$ be the set of all histories of infinite length, also called plays. Note that all the sets $\mathscr{H}_k$ have a measurable structure since they are products of measurable sets. Therefore, we can endow $\mathscr{H}_\infty$ with the σ-algebra generated by the cylinder sets. We also denote by $\mathscr{H} := \bigcup_{k \geqslant 1} \mathscr{H}_k$ the set of all finite histories.

A *(behavioral) strategy* of player MIN is a map $\sigma : \mathscr{H} \to \Delta(A)$ such that, for every finite history $h_k \in \mathscr{H}_k$ with $h_k = (s_1, a_1, b_1, \ldots, s_k)$, we have $\sigma(A_{s_k} \mid h_k) = 1$. We denote by $\mathcal{S}$ the set of all strategies of MIN. Similarly, the set $\mathcal{T}$ of strategies of player MAX is defined as the set of all maps $\tau : \mathscr{H} \to \Delta(B)$ such that $\tau(\cdot \mid h_k)$ is an element of $\Delta(B_{s_k})$ for every finite history h_k ending with s_k. A strategy is said to be *deterministic* or *pure* if the image of any finite history is a Dirac measure. In other words, a deterministic strategy of MIN (resp., MAX) can be identified with a map $\sigma : \mathscr{H} \to A$ (resp., $\tau : \mathscr{H} \to B$) such that $\sigma(h_k) \in A_{s_k}$ (resp., $\tau(h_k) \in B_{s_k}$) for every finite history h_k ending with s_k. We say that $\sigma(h_k)$ (resp., $\tau(h_k)$) is a *pure* action, as opposed to a *mixed* action.

Let us also introduce the following classes of strategies. A strategy is *Markovian* if it only depends on the length of the history and on its last state. A strategy σ of MIN (resp., τ of MAX) is *stationary* if it only depends on the current state, regardless of all the past information, that is, if there exists a map $\mu : S \to \Delta(A)$ (resp., $\nu : S \to \Delta(B)$) such that $\sigma(\cdot \mid h_k) = \mu(\cdot \mid s_k) \in \Delta(A_{s_k})$ (resp., $\tau(\cdot \mid h_k) = \nu(\cdot \mid s_k) \in \Delta(B_{s_k})$) for every finite history h_k ending with s_k.

Finally, an initial state $s \in S$ and a pair of strategies $(\sigma, \tau) \in \mathcal{S} \times \mathcal{T}$ define, by Kolmogorov extension theorem, a unique probability measure on the set of plays $\mathscr{H}_\infty$. We denote by $\mathbb{E}_{s,\sigma,\tau}$ the expectation with respect to that probability measure.

Global payoffs

Given a sequence of stage payments, generated by an initial state s and a choice of strategies $(\sigma, \tau) \in \mathcal{S} \times \mathcal{T}$, there are several ways of defining a global payoff. In this work, we focus our attention particularly on games played in finitely many stages. For a finite horizon $k \geqslant 1$, the payoff is thus given by the expected value of the sum of the k first stage payments, that is,

$$\gamma^k(s, \sigma, \tau) := \mathbb{E}_{s,\sigma,\tau}\left[\sum_{\ell=1}^{k} r(s_\ell, a_\ell, b_\ell)\right] .$$

Let us mention that a second classical way to evaluate the payoff is to fix a discount factor $\delta \in (0,1]$ and to consider the discounted sum of the stage payments:

$$\gamma^{\delta}(s,\sigma,\tau) := \mathbb{E}_{s,\sigma,\tau}\left[\delta \sum_{\ell=1}^{\infty}(1-\delta)^{(\ell-1)} r(s_\ell,a_\ell,b_\ell)\right] \ .$$

The methods used to deal with problems arising from these payoffs are similar. They usually boil down to consider strategy spaces as compact sets for some topology, under which the payoff function satisfies some continuity assumptions (the approach is usually referred to as the "compact case").

Another way to define the payoff is to consider the asymptotic behavior of the plays, on contrary to the finite-horizon or the discounted payoffs which depend essentially on a finite number of stages. Here, we shall especially consider the *limiting average reward criterion*, defined by

$$\gamma^{AR}(s,\sigma,\tau) := \liminf_{k\to\infty}\ \mathbb{E}_{s,\sigma,\tau}\left[\frac{1}{k}\sum_{\ell=1}^{k} r(s_\ell,a_\ell,b_\ell)\right] \ .$$

The problems based on this evaluation are also known as *ergodic control problems* in optimal control theory. Also, a stochastic game with this payoff is usually referred to as *mean-payoff game* in the computer science literature.

Value

Given a global payoff function γ, an initial state $s \in S$, and strategies $\sigma \in \mathcal{S}$ and $\tau \in \mathcal{T}$, player MAX intends to maximize the quantity $\gamma(s,\sigma,\tau)$, while player MIN wants to minimize it. The stochastic game Γ with payoff function γ has a *value* in $s \in S$ if

$$\sup_{\tau\in\mathcal{T}}\ \inf_{\sigma\in\mathcal{S}}\ \gamma(s,\sigma,\tau) = \inf_{\sigma\in\mathcal{S}}\ \sup_{\tau\in\mathcal{T}}\ \gamma(s,\sigma,\tau) \ .$$

In that case, the value, denoted by v_s, is equal to either of the terms in the above equality. Its existence means that for every $\varepsilon > 0$, player MAX can guarantee $v_s - \varepsilon$, that is, he has an ε-optimal strategy $\tau_\varepsilon \in \mathcal{T}$ such that

$$\forall \sigma \in \mathcal{S} \ , \quad \gamma(s,\sigma,\tau_\varepsilon) \geqslant v_s - \varepsilon \ ,$$

and dually player MIN can guarantee $v_s + \varepsilon$, that is, he has an ε-optimal strategy $\sigma_\varepsilon \in \mathcal{S}$ such that

$$\forall \tau \in \mathcal{T} \ , \quad \gamma(s,\sigma_\varepsilon,\tau) \leqslant v_s + \varepsilon \ .$$

If the *k-stage game* has a value in s, we denote it by v_s^k; if the *δ-discounted game* has a value in s, we denote it by v_s^δ; and if the *limiting average reward game*, also known as *mean payoff game*, has a value in s, we denote it by v_s^{AR}.

Once a payoff is fixed, a first problem is to find conditions under which the value exists. A second problem is to characterize the optimal, or ε-optimal, strategies. Moreover, for finite-horizon or discounted stochastic games, another classical problem concerns the asymptotic behavior of the value as the horizon goes to infinity or the discount factor vanishes. All these issues may be considered from either a theoretical or an algorithmic point of view.

In the finite-horizon case or the discounted case, the values are known to exist under standard assumptions, see [Sor02]. Let us mention the following two cases (recalling that in our framework, the state space is finite):

- the action spaces are finite (we say that the stochastic game Γ is *finite*);

– the sets of admissible actions are compact Hausdorff in every state, and the payment and transition functions are continuous.

Note that for particular classes of stochastic games, these values exist with mild assumptions. The class of perfect-information stochastic games, which will be presented in the next subsection, is one of them. Let us also mention the case of *Markov decision processes* (MDPs for short), which are one-player stochastic games, i.e., games in which one of the players has only a single admissible action in each state. Then, it suffices that the payoff function is bounded to ensure the existence of the value in finite horizon.

With the limiting average reward criterion, proving the existence of the value or finding optimal strategies are more complex tasks. However, the value is known to exist in the finite setup, that is, with finite state space and action spaces. In this framework, we have furthermore $v_s^{AR} = \lim_{\delta \to 0} v_s^\delta = \lim_{k \to \infty} v_s^k / k$. We refer the reader to [FV97] for a treatment of these issues in the case of stochastic games, and to [HLL96, HLL99] in the case of MDPs.

2.1.2 Games with perfect information

In this subsection, we introduce the subclass of *stochastic games with perfect information*. It differs from the general model in the structure of information: a stochastic game has perfect information when in each state one player has a unique admissible action, hence his decision has no effect on the payment nor the transition probabilities. If the player with trivial action space is always the same, the game is an MDP.

In this thesis, we consider a more general model where in each state, players take decisions in turn (MIN then MAX). Here, a perfect-information (zero-sum) stochastic game is defined by a 7-tuple $\Gamma := (S, A, B, K_A, K_B, r, p)$, where S is the state space, which is assumed to be finite, and A and B are the action spaces of players MIN and MAX respectively, which are assumed to be measurable sets. The constraint set of MIN is the same as the general model, that is, $K_A \subset S \times A$ is a measurable set such that each section $A_s := \{a \in A \mid (s, a) \in K_A\}$ represents the admissible actions of MIN in state s. As for the constraint set of MAX, it is a measurable set $K_B \subset S \times A \times B$, so that his admissible actions depend on the current state and the action chosen by the first player. If MIN has chosen action $a \in A_s$ in state s, then the set of admissible actions of MAX is

$$B_{s,a} := \{b \in B \mid (s, a, b) \in K_B\} \ .$$

We assume that all the sets of admissible actions are nonempty, and we let

$$K := \{(s, a, b) \mid s \in S,\ a \in A_s,\ b \in B_{s,a}\}$$

be the domain of definition of the payment function r and the transition function p, which are defined as in the general model.

The game is played repeatedly in discrete time like the "imperfect information" case, except that in stage $\ell \geqslant 1$, if the current state is s_ℓ, player MIN first select an action $a_\ell \in A_{s_\ell}$, then MAX observes this move and chooses an action $b_\ell \in B_{s_\ell, a_\ell}$.

To introduce the notion of strategy in our model, we need to distinguish histories of player MIN from histories of player MAX. The set of admissible histories of length $k \geqslant 1$ for player MIN is $\mathscr{H}_k := K^{k-1} \times S$, and we denote by $\mathscr{H} := \bigcup_{k \geqslant 1} \mathscr{H}_k$ the set of all finite histories of MIN. As for player MAX, the set of admissible histories of length $k \geqslant 1$ is $\mathscr{H}'_k := K^{k-1} \times K_A$, and we denote by $\mathscr{H}' := \bigcup_{k \geqslant 1} \mathscr{H}'_k$ the set of all finite histories for MAX.

A strategy of player MIN is a map $\sigma : \mathscr{H} \to \Delta(A)$ such that, for every finite history $h_k \in \mathscr{H}_k$ with $h_k = (s_1, a_1, b_1, \dots, s_k)$, we have $\sigma(\cdot \mid h_k) \in \Delta(A_{s_k})$ (and $\mathcal{S}$ denotes the set of all his strategies). As for player MAX, the set of his strategies, denoted by $\mathcal{T}'$, is composed of all the maps $\tau : \mathscr{H}' \to \Delta(B)$ such that $\tau(\cdot \mid h_k)$ is an element of $\Delta(B_{s_k, a_k})$ for every finite history h_k

ending with (s_k, a_k). The notion of Markovian, stationary, or deterministic strategy are defined in the same way as in the general case.

Given a payoff evaluation γ, the perfect-information game Γ has a value in state $s \in S$ if

$$\sup_{\tau \in \mathcal{T}'} \inf_{\sigma \in \mathcal{S}} \gamma(s, \sigma, \tau) = \inf_{\sigma \in \mathcal{S}} \sup_{\tau \in \mathcal{T}'} \gamma(s, \sigma, \tau) \ .$$

Contrary to the general case, the existence of the value requires less hypotheses, in particular regarding the topology of the action spaces and the regularity of the payment and transition functions. For instance, the existence is guaranteed in the finite-horizon case or the discounted case as soon as the payment function is bounded (with no continuity assumptions on r or p, nor topological assumptions on A or B). Furthermore, in the finite case, existence of deterministic stationary strategies was shown for discounted games [Sha53] and mean-payoff games [LL69].

In the sequel, the same notation Γ will refer either to a stochastic game (with imperfect information) or a perfect-information stochastic game. When no precision is made, a stochastic game will always refer to the general model.

2.1.3 Asymptotic approach: the mean payoff

A major topic in the theory of repeated games concerns the asymptotic behavior of the value in finite horizon when the horizon goes to infinity. Particularly, given an initial state $s \in S$, a standard question is to understand when the sequence of mean values per time unit, $(v_s^k/k)_{k \geqslant 1}$, has a finite limit. This limit, when it exists, is referred to as the *mean payoff* (in state s), and we denote it by χ_s:

$$\chi_s := \lim_{k \to \infty} \frac{v_s^k}{k} \ , \quad s \in S \ .$$

We shall denote by χ the *mean payoff vector*, i.e., the vector in $\mathbb{R}^S$ whose entries are χ_s with $s \in S$.

The latter question was first addressed for particular classes of repeated games with finite state space. Everett [Eve57] proved the existence of the mean payoff for *recursive games*, a class of stochastic games where a nonzero payment only occurs in *absorbing states* (states that cannot be escaped from and where the payoff remains the same). Kohlberg [Koh74] proved it for *absorbing games*, where the transition are deterministic and some state are absorbing. Then, Bewley and Kohlberg [BK76] showed that the limit exists for all finite stochastic games. For more recent results, see [RS01a, Sor04] or [BGV15]. In the last reference, Bolte, Gaubert and Vigeral proved that the mean payoff exists for *definable stochastic games*, where all the data (action spaces, payment and transition functions) are definable in some o-minimal structure, a condition that often holds in practice. In the one-player case, the existence of the mean payoff is ensured with even milder conditions. Thus, Renault [Ren11] proved that for any MDP with bounded payment function (and finite state space) the mean payoff exists – actually, the result is stronger and concerns the existence of the *uniform value*. Note that existence results concerning the asymptotic approach were also obtained for games with incomplete information [AM95, MZ72].

On the other hand, the sequence of mean values may not converge, even with standard hypotheses. Vigeral [Vig13] provided an example of a stochastic game with finite state space, compact action spaces, and continuous payment and transition functions, such that the mean value in finite horizon does not converge as the horizon goes to infinity. See also [Zil16] for similar nonexistence results.

2.2 Operator approach of stochastic games

2.2.1 Shapley operator and recursive structure

General case

Let Γ be a stochastic game with state space $S = [n]$. To any vector $x \in \mathbb{R}^n$ and any state $i \in [n]$, we associate an auxiliary "one-shot" game, i.e., a repeated game played in one stage, defined by the same action spaces and constraint sets as Γ and by the payment function

$$(a, b) \mapsto r(i, a, b) + \sum_{j \in [n]} x_j \, p(j \mid i, a, b) \quad , \quad \forall (a, b) \in A_i \times B_i \ .$$

Assume that the latter one-stage game is well defined and has a value, meaning that

$$\inf_{\mu \in \Delta(A_i)} \sup_{\nu \in \Delta(B_i)} \int_{A_i} \int_{B_i} \Big(r(i, a, b) + \sum_{j \in [n]} x_j \, p(j \mid i, a, b) \Big) \, d\nu(b) \, d\mu(a) =$$
$$\sup_{\nu \in \Delta(B_i)} \inf_{\mu \in \Delta(A_i)} \int_{B_i} \int_{A_i} \Big(r(i, a, b) + \sum_{j \in [n]} x_j \, p(j \mid i, a, b) \Big) \, d\mu(a) \, d\nu(b) \ .$$

We denote by $T_i(x)$ this value, which represents the value of the stochastic game Γ starting in state i, played in one stage, and with an additional payoff of x_j if the terminal state is j. Let us mention that this value exists, according to Sion's minimax theorem [Sio58], as soon as the sets of actions are compact Hausdorff and the payment and transition functions are continuous. Then, we define the *Shapley operator* $T : \mathbb{R}^n \to \mathbb{R}^n$ of the stochastic game Γ as the operator whose ith coordinate map is given by

$$\begin{aligned} T_i(x) &= \inf_{\mu \in \Delta(A_i)} \sup_{\nu \in \Delta(B_i)} \Big(r(i, \mu, \nu) + \sum_{j \in [n]} x_j \, p(j \mid i, \mu, \nu) \Big) \\ &= \sup_{\nu \in \Delta(B_i)} \inf_{\mu \in \Delta(A_i)} \Big(r(i, \mu, \nu) + \sum_{j \in [n]} x_j \, p(j \mid i, \mu, \nu) \Big) \quad , \quad \forall x \in \mathbb{R}^n \ , \end{aligned} \tag{2.1}$$

where, given two probability measures $\mu \in \Delta(A_i)$ and $\nu \in \Delta(B_i)$,

$$f(i, \mu, \nu) := \int_{A_i} \int_{B_i} f(i, a, b) \, d\nu(b) \, d\mu(a)$$

is the multilinear extension of any function $f : K \to \mathbb{R}$.

This operator encompasses the recursive structure of Γ, as shown by Shapley [Sha53] in the case of finite stochastic games. Indeed, if $v^k := (v_i^k)_{i \in [n]}$ denotes the *value vector* of the k-stage game Γ, i.e., the vector in $\mathbb{R}^n$ the ith entry of which is the value of the game in finite horizon k with initial state i, then by applying a dynamic programming principle, we have

$$v^0 = 0 \qquad \text{and} \qquad v^{k+1} = T(v^k) \quad , \quad \forall k \in \mathbb{N} \ . \tag{2.2}$$

As a consequence, if the mean payoff vector exists, then it is given by the growth rate of the orbit of T with initial point 0:

$$\chi = \lim_{k \to \infty} \frac{T^k(0)}{k} \ , \tag{2.3}$$

where $T^k := T \circ \cdots \circ T$ denotes the kth iterate of T. The "operator approach", initiated by Rosenberg and Sorin [RS01a], consists in studying the properties of the Shapley operator and its iterates to infer results about the asymptotic behavior of the sequence of values $(v^k)_{k \in \mathbb{N}}$.

Perfect-information case

When Γ is a perfect-information stochastic game with state space $S = [n]$, the Shapley operator simplifies. The fact that players choose actions one after the other with a perfect knowledge of all available information makes the use of mixed actions (i.e., choices of probability measures over the set of admissible actions) irrelevant in finitely repeated games. Thus, with perfect information, we only need to consider pure actions, and the auxiliary one-shot game with initial state i and terminal payoff $x \in \mathbb{R}^n$ has a value as soon as

$$\inf_{a \in A_i} \sup_{b \in B_{i,a}} \Big(r(i,a,b) + \sum_{j \in [n]} x_j \, p(j \mid i,a,b) \Big) \in \mathbb{R} \enspace .$$

In particular, this is the case when the payment function is bounded (the action spaces being any sets and no continuity assumption on the payment and transition functions being made). Then, the Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ has a simpler form, its ith coordinate writing

$$T_i(x) = \inf_{a \in A_i} \sup_{b \in B_{i,a}} \Big(r(i,a,b) + \sum_{j \in [n]} x_j \, p(j \mid i,a,b) \Big) \enspace , \quad \forall x \in \mathbb{R}^n \enspace . \tag{2.4}$$

Note that the recursive property (2.2) of the values in finite horizon still holds.

2.2.2 Axiomatization of Shapley operators

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of a stochastic game with either perfect information (2.4) or imperfect information (2.1). We endow $\mathbb{R}^n$ with the usual partial order and denote by e the unit vector, i.e., the vector whose entries are all equal to 1. Then, it is readily seen that T satisfies the following two properties:

- *monotonicity*:
$$x \leqslant y \implies T(x) \leqslant T(y) \enspace , \quad \forall x, y \in \mathbb{R}^n \enspace ; \tag{M}$$
- *additive homogeneity*:
$$T(x + \lambda e) = T(x) + \lambda e \enspace , \quad \forall (\lambda, x) \in \mathbb{R} \times \mathbb{R}^n \enspace . \tag{AH}$$

An immediate consequence is that T is *sup-norm nonexpansive*:

$$\| T(x) - T(y) \|_\infty \leqslant \| x - y \|_\infty \enspace , \quad \forall x, y \in \mathbb{R}^n \enspace , \tag{N$_\infty$}$$

where $\|z\|_\infty := \max\{|z_i| \mid 1 \leqslant i \leqslant n\}$ is the standard supremum norm in $\mathbb{R}^n$. Let us mention that Crandall and Tartar [CT80] proved that under (AH), the properties (M) and (N$_\infty$) are equivalent.

Remark 2.1. The nonexpansiveness of T implies that the asymptotic behavior of the sequence $(T^k(x)/k)_{k \geqslant 1}$ is independent of the vector $x \in \mathbb{R}^n$: either it always converges (to a finite limit), or it never converges. Thus, the mean payoff vector (2.3) is given by the growth rate of *any* orbit of T.

The importance of monotonicity and additive homogeneity was recognized early on in dynamic programming [Bla65]. The study of operators satisfying these properties, so called operator approach, has proven useful not only in game theory [BK76, FV97, RS01a, Sor04, Vig10, BGV15], but also in optimal control [Koh80, Whi83, Kol92], or in the modelling of discrete event systems (such as communication networks, digital circuits, manufacturing processes) [BCOQ92, Gun03].

Conversely, Kolokoltsov showed that every operator from $\mathbb{R}^n$ to itself which is monotone and additively homogeneous can be written as the Shapley operator (2.1) of a stochastic game.

Theorem 2.2 ([Kol92], see also [KM97, Th. 2.15])**.** *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a monotone and additively homogeneous map. Then, there exist a payment function $r : [n] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and a transition function p from $[n] \times \mathbb{R}^n \times \mathbb{R}^n$ to the standard simplex of $\mathbb{R}^n$, such that the ith coordinate map of F is given by*

$$\begin{aligned} F_i(x) &= \min_{a \in \mathbb{R}^n} \max_{b \in \mathbb{R}^n} \Big(r(i,a,b) + \sum_{j \in [n]} x_j \, p_j(i,a,b) \Big) \\ &= \max_{b \in \mathbb{R}^n} \min_{a \in \mathbb{R}^n} \Big(r(i,a,b) + \sum_{j \in [n]} x_j \, p_j(i,a,b) \Big) \enspace , \quad \forall x \in \mathbb{R}^n \enspace . \end{aligned}$$

The proof of this result, which only holds in finite dimension, is based on a minimax representation formula due to Evans, which applies more generally to Lipschitz functions [Eva84]. Rubinov and Singer proposed a similar representation, but with perfect information. Moreover, in their formula, the transition probabilities are degenerate, i.e., deterministic.

Theorem 2.3 ([RS01b, Th. 5.3])**.** *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a monotone and additively homogeneous map. Then, its ith coordinate map is given by*

$$F_i(x) = \min_{y \in \mathbb{R}^n} \max_{j \in [n]} \big(x_j - y_j + F_i(y) \big) \enspace , \quad \forall x \in \mathbb{R}^n \enspace .$$

Note that Gunawardena and Sparrow obtained independently an equivalent result [Gun03, Prop. 2.3]. Also note that in the context of abstract convexity, a representation of functions on cones with similar properties has been given in [DMLR04].

The above considerations motivate us to identify, in this thesis, Shapley operators with maps that satisfy properties (M) and (AH).

Definition 2.4. We call *(abstract) Shapley operator* on $\mathbb{R}^n$ any map from $\mathbb{R}^n$ to itself that is both monotone (M) and additively homogeneous (AH).

2.2.3 Ergodic equation

As mentioned above, the operator approach for stochastic games consists in studying the properties of Shapley operators in order to infer information about the asymptotic behavior of the values in finite horizon. A basic tool in this respect is the following equation, called *ergodic equation* (also known as the *average reward optimality equation* in the optimal control and MDP literature):

$$T(u) = \lambda e + u \enspace , \quad (\lambda, u) \in \mathbb{R} \times \mathbb{R}^n \enspace . \tag{2.5}$$

We say that the ergodic equation is solvable if their exist both a scalar $\lambda \in \mathbb{R}$ and a vector $u \in \mathbb{R}^n$ for which (2.5) holds. If this is the case, we shall say that λ is the *ergodic constant* and u is a *bias vector* of T. Since the solvability of the ergodic equation can be seen as a nonlinear spectral problem in the additive framework (see Section 2.3), we shall also call λ the *(additive) eigenvalue* of T, and u an *(additive) eigenvector*. Consequently, saying that T has an (additive) eigenvalue is equivalent to saying that the ergodic equation is solvable for T.

Observe that if the ergodic equation is solvable, then $T^k(u) = k\lambda e + u$ for all integers $k \in \mathbb{N}$, and so the mean payoff vector exists (see Remark 2.1) and each of its components is equal to λ, that is, the mean payoff of the associated game is independent of the initial state. It follows that the ergodic constant, when it exists, is necessarily unique. Also, since T is additively homogeneous, any vector $u + \alpha e$ with $\alpha \in \mathbb{R}$ is a bias vector. We shall say that u is defined "up to an additive constant". Here, by "constant (vector)", we mean a vector proportional to the unit vector e.

The ergodic equation is useful not only for the study of the mean-payoff vector, but also for the study of the limiting average reward game. In particular, knowing explicitly an eigenvector gives access to optimal or ε-optimal stationary strategies, see [KY92] for the deterministic case, which readily extends to the stochastic case. See also [FV97] and [HLL96, HLL99] for the links between the ergodic equation and optimal strategies in the framework of stochastic games and MDPs, respectively.

The main purpose of this work is to find conditions guaranteeing the solvability of the ergodic equation, as well as to characterize the set of bias vectors.

2.2.4 Piecewise affine operators and invariant half-lines

We conclude this section by mentioning the particular case of perfect-information finite stochastic games. Recall that a stochastic game (with perfect information or not) is finite if the state space and the action spaces are finite. In that case, the Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is *piecewise affine* , meaning that $\mathbb{R}^n$ can be covered by finitely many polyhedra (a polyhedron being an intersection of finitely many half-spaces) such that the restriction of the coordinate functions of T to each polyhedron are affine.

Note that continuous piecewise affine functions are exactly the ones that can be represented as a minimax over finite sets of affine functions, i.e., as in (2.4) where the infimum and the supremum are taken over finite sets and the coefficients can be any real numbers, see [Ovc02, AT07].

Then, the following result applies.

Theorem 2.5 ([Koh80, Th. 2.1]). *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a piecewise affine map that is nonexpansive in any norm. Then, it has an* invariant half-line, *that is, there exist two vectors $u, \nu \in \mathbb{R}^n$ such that*

$$F(u + \alpha\nu) = u + (\alpha + 1)\nu$$

for all $\alpha \geqslant 0$. Moreover, the vector ν is unique.

Hence, the Shapley operator T of a perfect-information finite stochastic game has an invariant half-line, and so there exist vectors u, ν as in the above theorem. This yields $T^k(u) = u + k\nu$ for all integers $k \in \mathbb{N}$, which implies that the mean-payoff vector exists (see Remark 2.1) and is equal to ν. Moreover, it is easily seen that the solvability of the ergodic equation (2.5) is equivalent to the mean payoff being a constant vector.

2.3 Connection with nonlinear Perron-Frobenius theory

2.3.1 Additive versus multiplicative framework

Nonlinear Perron-Frobenius theory is primarily concerned with the study of nonexpansive maps f acting on a closed convex cones K in a Banach space. An important special case arises when f is homogeneous of degree one and preserves the partial order induced by K. The central question is to know if f has an eigenvector in the interior of K, i.e., if there exists a vector $u \in \operatorname{int}(K)$ such that

$$f(u) = \lambda u \tag{2.6}$$

for some scalar $\lambda > 0$. Other standard problems include the characterization of the eigenvalue, which may be seen as a spectral radius, or the uniqueness, up to the product by a positive scalar, of the eigenvector, see [LN12].

These problems are already interesting when K is the standard nonnegative cone of $\mathbb{R}^n$, namely $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n \mid x \geqslant 0\}$, having applications, e.g., in matrix scaling problems or nonnegative tensor problems. Note that if f is a $n \times n$ nonnegative matrix, we recover the classical Perron-Frobenius theory.

In this "multiplicative" framework, the aforementioned problems have been considered in particular by Nussbaum [Nus88, Nus89]. But when f leaves the interior of $\mathbb{R}_+^n$ invariant, one can equivalently work in the "additive" framework, see [GG04]. Indeed, the space $\mathbb{R}^n$ can be placed in bijection with the interior of $\mathbb{R}_+^n$ via the map $\exp : \mathbb{R}^n \to \operatorname{int}(\mathbb{R}_+^n)$ and its inverse $\log : \operatorname{int}(\mathbb{R}_+^n) \to \mathbb{R}^n$, which applies respectively the exponential and the logarithm componentwise. Hence, any map $f : \operatorname{int}(\mathbb{R}_+^n) \to \operatorname{int}(\mathbb{R}_+^n)$ which is homogeneous of degree one and preserves the standard partial order of $\mathbb{R}^n$ is conjugated with a unique map $T : \mathbb{R}^n \to \mathbb{R}^n$, namely $T := \log \circ f \circ \exp$, which is monotone (M) and additively homogeneous (AH), i.e., a Shapley operator. Then, through "$\log$-glasses", it is readily seen that the nonlinear eigenproblem (2.6) is equivalent to finding a solution to the ergodic equation (2.5).

Thus, it is equivalent to work in the multiplicative framework and the additive one, as far as the above properties are concerned, and all the results in the present work, established in the additive framework, can be translated in the multiplicative one.

2.3.2 Hilbert's seminorm and invariant sets

A useful tool in nonlinear Perron-Frobenius theory is *Hilbert's seminorm*. In the space $\mathbb{R}^n$, it is defined by

$$\|x\|_{\mathrm{H}} := \max_{i \in [n]} x_i - \min_{j \in [n]} x_j = \max_{i,j \in [n]} (x_i - x_j) \ , \quad \forall x \in \mathbb{R}^n \ .$$

The following result justifies the interest of Hilbert's seminorm for the nonlinear spectral problem we are considering. As mentioned in [GG04], it can be seen as a special case, in the additive framework, of a general result by Nussbaum [Nus88, Th. 4.1].

Theorem 2.6 ([GG04, Th. 9]). *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a Shapley operator. The ergodic equation (2.5) is solvable if, and only if, the orbit $\{T^k(x) \mid k \in \mathbb{N}\}$ is bounded in Hilbert's seminorm for some (hence all) $x \in \mathbb{R}^n$.*

The above theorem provides a method to tackle the problem of the solvability of the ergodic equation, method which consists in finding invariant subsets that are bounded in Hilbert's seminorm. Following this direction, let us introduce three particular families of subsets that are structurally invariant under all Shapley operators T. Given scalars $\alpha, \beta \in \mathbb{R}$, we define:

– the *sub-eigenspace*

$$\mathcal{S}_\alpha(T) := \{x \in \mathbb{R}^n \mid \alpha e + x \leqslant T(x)\} \ ;$$

– the *super-eigenspace*

$$\mathcal{S}^\beta(T) := \{x \in \mathbb{R}^n \mid T(x) \leqslant \beta e + x\} \ ;$$

– the *slice space*

$$\mathcal{S}_\alpha^\beta(T) := \mathcal{S}_\alpha(T) \cap \mathcal{S}^\beta(T) = \{x \in \mathbb{R}^n \mid \alpha e + x \leqslant T(x) \leqslant \beta e + x\} \ .$$

Even without the boundedness property, the nonemptiness of those spaces may be useful. Indeed, a pair $(\alpha, x) \in \mathbb{R} \times \mathbb{R}^n$ such that $x \in \mathcal{S}_\alpha(T)$, or $(\beta, x) \in \mathbb{R} \times \mathbb{R}^n$ such that $x \in \mathcal{S}^\beta(T)$, may be seen as the solution of a generalized ergodic equation. It provides information about the asymptotic behavior of the iterates of T, as shown by the following result. Note that in the MDP literature, this tool is known as the *average-cost optimality inequality*, see [HLL96].

Proposition 2.7 (compare with [Sor04, Prop. 7])**.** *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a Shapley operator. Assume that the sub-eigenspace $\mathcal{S}_\alpha(T)$ is nonempty for a given scalar α. Then*

$$\liminf_{k\to\infty} \frac{T^k(0)}{k} \geqslant \alpha e \enspace .$$

Similarly, if the super-eigenspace $\mathcal{S}^\beta(T)$ is not empty for $\beta \in \mathbb{R}$, then

$$\limsup_{k\to\infty} \frac{T^k(0)}{k} \leqslant \beta e \enspace .$$

Finally, let us mention that the slice space $\mathcal{S}_\alpha^\beta(T)$ is nonempty if $\alpha \leqslant \min_{i\in[n]} T_i(0)$ and $\beta \geqslant \max_{i\in[n]} T_i(0)$. Hence, the difficulty in applying Theorem 2.6 with some slice space resides in checking its boundedness with respect to Hilbert's seminorm.

2.3.3 Conditions for the solvability of the ergodic equation

The question of characterizing the Shapley operators for which one of the aforementioned invariant spaces is bounded in Hilbert's seminorm (but not necessarily all of them) is delicate, and depends heavily not only on the structure but also on the parameters defining the operator. On the other hand, there exist conditions ensuring the boundedness of *all* invariant spaces of a given family. These conditions, that we present in the sequel, rely on the asymptotic behavior of the Shapley operators. Hence, they are insensitive to perturbations that preserve the monotonicity and the additive homogeneity and that do not change the behavior at infinity.

Boundedness of super-eigenspaces

The first condition involves the super-eigenspaces (note that a dual statement holds for sub-eigenspaces). We need the following construction. To any Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$, we associate the directed graph $\mathcal{G}(T)$, with set of nodes $[n]$ and an arc from i to j if $T_i(\alpha e_{\{j\}})$ tends to $+\infty$ as α goes to $+\infty$, where $e_{\{j\}}$ is the jth vector of the canonical basis of $\mathbb{R}^n$. The directed graph $\mathcal{G}(T)$ is said to be strongly connected if there is a path between any two distinct nodes.

Theorem 2.8 ([GG04, Th. 10])**.** *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a Shapley operator. If $\mathcal{G}(T)$ is strongly connected, then all super-eigenspaces are bounded in Hilbert's seminorm.*

Note that the above condition is only sufficient, but in the same paper [GG04], a combinatorial characterization of the boundedness of all super-eigenspaces is given, involving aggregated graphs constructed in the same spirit as $\mathcal{G}(T)$.

Boundedness of slice spaces

The second condition involves the slice spaces. Let us denote by $\widehat{T}$ the *recession operator* of any Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$, defined by

$$\widehat{T}(x) := \lim_{\alpha\to+\infty} \frac{T(\alpha x)}{\alpha} \enspace , \quad \forall x \in \mathbb{R}^n \enspace . \tag{2.7}$$

Note that the recession operator does not always exist, but when it does it inherits from T the monotonicity and the additive homogeneity. In addition, it is *positively homogeneous*, meaning that

$$\widehat{T}(\alpha x) = \alpha \widehat{T}(x) \enspace , \quad \forall \alpha > 0 \enspace , \quad \forall x \in \mathbb{R}^n \enspace . \tag{PH}$$

As a consequence, any constant vector is a fixed point of $\widehat{T}$. We shall call such fixed points *trivial fixed points*.

Theorem 2.9 ([GG04, Th. 13])**.** *If a Shapley operator T on $\mathbb{R}^n$ has a recession operator whose only fixed points are trivial, then all slice spaces of T are bounded in Hilbert's seminorm.*

The above result provides a sufficient condition for the boundedness of all slice spaces, but a counterexample in [GG04] showed that the condition is not necessary. In particular, the following problem remained open.

Problem ([GG04])**.** Give a combinatorial condition or graph-theoretic characterization of the property that all slice spaces are bounded in Hilbert's seminorm.

The present work gives an answer to that problem.

Weakly convex operators

We mention that Cavazos-Cadena and Hernández-Hernández proved in [CCHH10], under a weak form of convexity, that $\widehat{T}$ has only trivial fixed points if, and only if, the ergodic equation is solvable for all operators $g + T$ with $g \in \mathbb{R}^n$. The weak convexity property was motivated by risk sensitive control problems. A typical example is the following Shapley operator T:

$$T_i(x) = \log \Big(\sum_{j \in [n]} M_{ij} \exp(x_j) \Big) \ , \quad i \in [n] \ , \ \ x \in \mathbb{R}^n \ , \tag{2.8}$$

where $M = (M_{ij})$ is a nonnegative matrix without zero row. Note that a supremum of weakly convex Shapley operators is weakly convex, hence, one can construct further examples of weakly convex operators by taking suprema of operators of the form (2.8).

This result gives a characterization, under the weakly convex property, of the boundedness of all slice spaces in Hilbert's seminorm. Indeed, it follows from Theorem 2.6 that all Shapley operators $g + T$ with $g \in \mathbb{R}^n$ have an eigenvector if all the slice spaces of T are bounded in Hilbert's seminorm.

Let us also mention that equivalent combinatorial conditions were given in the same paper, involving the directed graph $\mathcal{G}(T)$.

2.3.4 From nonlinear spectral problems to fixed point problems

The existence results in nonlinear Perron-Frobenius theory usually rely on classical fixed-point theorems. In this subsection, we detail the link between the two topics in regard to our problems.

Let $\mathbb{TP}^n$ be the space $\mathbb{R}^n$ quotiented by the equivalence relation $\sim$ defined as follows: for two vectors $x, y \in \mathbb{R}^n$, we write $x \sim y$ if there is a scalar $\alpha \in \mathbb{R}$ such that $x - y = \alpha e$. By analogy with the real projective space, we call it the *additive projective space*. We denote by $[x]$ the equivalence class of any vector $x \in \mathbb{R}^n$, which is the affine line directed by the unit vector e and passing through x. Observe that $\|x\|_{\mathrm{H}} = \|y\|_{\mathrm{H}}$ if $x \sim y$. Hence, Hilbert's seminorm can be quotiented into a norm of $\mathbb{TP}^n$, that we denote by q_{H}:

$$q_{\mathrm{H}}([x]) := \|x\|_{\mathrm{H}} \ , \quad \forall x \in \mathbb{R}^n \ .$$

Furthermore, $\mathbb{TP}^n$ has the structure of a vector space:

$$\alpha[x] + [y] := [\alpha x + y] \ , \quad \forall x, y \in \mathbb{R}^n \ , \quad \forall \alpha \in \mathbb{R} \ .$$

Thus, $(\mathbb{TP}^n, q_{\mathrm{H}})$ is a normed vector space of dimension $n - 1$.

Since any Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is additively homogeneous, it is readily seen that it can be quotiented into a map $[T] : \mathbb{TP}^n \to \mathbb{TP}^n$, sending each equivalence class $[x]$ to the class

$[T(x)]$. Moreover, it is a standard result that any Shapley operator is nonexpansive with respect to Hilbert's seminorm (see [GG04]), that is,

$$\|T(x) - T(y)\|_{\mathrm{H}} \leqslant \|x - y\|_{\mathrm{H}} \ , \quad \forall x, y \in \mathbb{R}^n \ . \tag{N$_\mathrm{H}$}$$

Hence, the quotiented map $[T]$ is nonexpansive with respect to the norm q_{H} on $\mathbb{TP}^n$.

Now observe that a vector $u \in \mathbb{R}^n$ is an eigenvector of T if, and only if, its equivalence class $[x] \in \mathbb{TP}^n$ is a fixed point of $[T]$. Thus, the ergodic problem (2.5) reduces to a fixed point problem involving a nonexpansive map.

CHAPTER 3

Minimax representation of positively homogeneous nonexpansive functions

3.1 Introduction

Several minimax representation theorems are available for monotone additively homogeneous operators on $\mathbb{R}^n$: characterization in terms of stochastic games [Kol92] (Theorem 2.2), or in terms of perfect-information repeated games with deterministic transitions [RS01b, Gun03] (Theorem 2.3). In the context of abstract convexity, we also refer to [DMLR04, DMLR08] for a characterization of similar operators in the multiplicative framework. The Fenchel-Legendre duality provides the same kind of representation theorems (but with no infimum) for operators, or real functions, that are in addition convex. It has been applied in infinite dimension to the case of convex risk measures [FS02] for instance.

An important subclass of monotone and additively homogeneous operators arises when considering positive homogeneity (PH). This is particularly the case for recession operators, used by Gaubert and Gunawardena [GG04] in nonlinear Perron-Frobenius theory, or for coherent risk measures [ADEH99, Del02] in infinite dimension.

Observe that if $T : \mathbb{R}^n \to \mathbb{R}^n$ is the Shapley operator of a stochastic game (with perfect information or not) $\Gamma(r) = ([n], A, B, K_A, K_B, r, p)$ the payment function of which is bounded, then its recession operator $\widehat{T}$ exists and it is the Shapley operator of the stochastic game $\Gamma(0)$ which has a zero payment function. However, the converse cannot be deduced from the representation theorems mentioned above: if $T : \mathbb{R}^n \to \mathbb{R}^n$ is an abstract Shapley operator that is also positively homogeneous, we cannot derive from Theorem 2.2 or 2.3 a representation formula of T as the dynamic programming operator of a game with zero payment function. In this chapter, we provide such a characterization. This leads us to the following definition.

Definition 3.1. We call *(abstract) payment-free Shapley operator* on $\mathbb{R}^n$ any self-map of $\mathbb{R}^n$ that is monotone (M), additively homogeneous (AH) and positively homogeneous (PH).

It is known that the two properties characterizing Shapley operators, (M) and (AH), are equivalent to nonexpansiveness with respect to some weak Minkowski norm. The latter are substitutes of norms that are not necessarily symmetric or separating. They have been studied in metric geometry, in particular by Papadopoulos and Troyanov [PT14], and arise naturally in the study of Shapley operators, see [GV12].

To get a minimax characterization of monotone, additively and positively homogeneous operators, we first establish a general minimax representation theorem which applies to real functions on a topological vector space that are nonexpansive with respect to any weak Minkowski norm. We then characterize nonexpansive maps that are positively homogeneous. As a corollary, we arrive at our main application: a representation theorem for dynamic programming operators of stochastic games with zero payments. This representation also leads to an approximation of the latter operators by polyhedral maps. Such approximation can been used in the setting of "max-plus basis methods" for the numerical solutions of Hamilton-Jacobi type partial differential equations, see [McE06, AGL08] for background. We also arrive at a representation theorem for nonconvex risk measures.

The results persented in this chapter are based on the preprint [AGH16].

3.2 Preliminaries: nonexpansiveness and weak Minkowski norms

In this chapter, we denote by $\mathcal{V}$ a real topological vector space (t.v.s. for short). We denote by $\mathcal{V}^*$ the dual space of $\mathcal{V}$ (i.e., the space of linear forms on $\mathcal{V}$), by $\mathcal{V}'$ its topological dual space (i.e., the space of continuous linear forms on $\mathcal{V}$), and by $\langle \cdot, \cdot \rangle$ the duality product.

We shall specially consider the situation in which $\mathcal{V}$ is a vector space with an (Archimedean) order unit. This means that we assume that $\mathcal{V}$ is a real vector space with an order relation, denoted by $\leqslant$, that is compatible with the algebraic structure of $\mathcal{V}$, that is, satisfying the following two axioms:

(i) $x \leqslant y \implies x + z \leqslant y + z \ , \quad \forall z \in \mathcal{V}$;

(ii) $x \leqslant y \implies \lambda x \leqslant \lambda y \ , \quad \forall \lambda \in \mathbb{R}_+$;

where $\mathbb{R}_+$ is the set of nonnegative real numbers. We also assume that $\mathcal{V}$ is equipped with a special vector denoted by $e_{\mathcal{V}}$ (or simply by e if the context is clear), called an *order unit* and such that for every $x \in \mathcal{V}$, there exists a scalar $\lambda > 0$ satisfying $x \leqslant \lambda e_{\mathcal{V}}$. Finally, we assume that this order unit is *Archimedean*, i.e., for any $x \in \mathcal{V}$, we have $x \geqslant 0$ if $\lambda e_{\mathcal{V}} + x \geqslant 0$ for all scalars $\lambda > 0$. Such a space will be endowed with the topology defined by the following norm:

$$\|x\|_{e_{\mathcal{V}}} = \inf\{\lambda \in \mathbb{R} \mid -\lambda e_{\mathcal{V}} \leqslant x \leqslant \lambda e_{\mathcal{V}}\} \ , \quad x \in \mathcal{V} \ , \tag{3.1}$$

see [PT09]. Note that if $\mathcal{V}$ is the Euclidean space $\mathbb{R}^n$ equipped with the standard partial order, then the standard unit vector of $\mathbb{R}^n$ is an Archimedean order unit, and the corresponding norm (3.1) is the usual supremum norm. Hence, we will abusively refer to (3.1) as the "supnorm" even in general situations.

An important particular case is obtained when $\mathcal{V}$ is an AM-space with unit, i.e., a Banach lattice equipped with an order unit and such that the norm satisfies (3.1), see [AB06]. By the Kakutani-Krein theorem, any AM-space with unit is isomorphic (lattice isometric) to a space $\mathcal{C}(K)$ of continuous functions over some compact Hausdorff set K, equipped with the supnorm, see [AB06, Th. 8.29] or [Sch74, Ch. II, Th. 7.4].

Given two vector spaces with an order unit, $(\mathcal{V}, e_{\mathcal{V}})$ and $(\mathcal{W}, e_{\mathcal{W}})$, we will be interested in maps $F : \mathcal{V} \to \mathcal{W}$ that satisfy, for all $x, y \in \mathcal{V}$, some of the following properties:

– monotonicity:

$$x \leqslant y \implies F(x) \leqslant F(y) \ ; \tag{M}$$

– additive homogeneity:

$$F(\lambda e_{\mathcal{V}} + x) = \lambda e_{\mathcal{W}} + F(x) \ , \quad \forall \lambda \in \mathbb{R} \ ; \tag{AH}$$

– additive subhomogeneity:

$$F(\lambda e_{\mathcal{V}} + x) \leqslant \lambda e_{\mathcal{W}} + F(x) \ , \quad \forall \lambda > 0 \ ; \tag{ASH}$$

– positive homogeneity:

$$F(\lambda x) = \lambda F(x) \ , \quad \forall \lambda > 0 \ ; \tag{PH}$$

– sup-norm nonexpansiveness:

$$\|F(x) - F(y)\|_{e_{\mathcal{W}}} \leqslant \|x - y\|_{e_{\mathcal{V}}} \ . \tag{N_∞}$$

The importance of monotonicity and additive homogeneity in optimal control and game theory has been pointed in Subsection 2.2.2. Crandall and Tartar [CT80] showed that

$$\text{(M) and (AH)} \iff (N_\infty) \text{ and (AH)} \ ,$$

when $\mathcal{V} = \mathcal{W}$ is a L^∞ space. It is also known that

$$\text{(M) and (ASH)} \iff \text{(M) and } (N_\infty) \ ,$$

see, e.g., [AG03]. These relations are readily generalized to any vector spaces with an order unit.

The monotonicity and additive homogeneity properties turn out to be related with nonexpansiveness with respect to weak Minkowski norms. By *weak Minkowski norm* on $\mathcal{V}$, we mean a function $q : \mathcal{V} \to \mathbb{R}$ that is subadditive (i.e., $q(x + y) \leqslant q(x) + q(y)$ for all $x, y \in \mathcal{V}$) and positively homogeneous (PH), hence convex but not necessarily symmetric (i.e., we do not require that $q(x) = q(-x)$ for all $x \in \mathcal{V}$). Our definition is a variant of the one in [PT14], where q is also required to be nonnegative and may take infinite values. We say that a real function on $\mathcal{V}$, $f : \mathcal{V} \to \mathbb{R}$, is nonexpansive with respect to q if $f(x) - f(y) \leqslant q(x - y)$ for all $x, y \in \mathcal{V}$.

When $\mathcal{V} = \mathbb{R}^n$, a useful example of weak Minkowski norm, arising in Hilbert geometry [PT14], is the "top" map, $\mathsf{t} : \mathbb{R}^n \to \mathbb{R}$, defined by

$$\mathsf{t}(x) := \max_{i \in [n]} x_i \ , \quad x \in \mathcal{V} \ ,$$

or its variant, $\mathsf{t}^+ : \mathbb{R}^n \to \mathbb{R}_+$, defined by

$$\mathsf{t}^+(x) := \max\{\mathsf{t}(x), 0\} \ , \quad x \in \mathcal{V} \ .$$

When $\mathcal{V} = \mathcal{W} = \mathbb{R}^n$, we shall consider for any map $F : \mathbb{R}^n \to \mathbb{R}^n$ the following properties:

– nonexpansiveness in t:

$$\mathsf{t}(F(x) - F(y)) \leqslant \mathsf{t}(x - y) \ ; \tag{N_{t}}$$

– nonexpansiveness in t^+:

$$\mathsf{t}^+(F(x) - F(y)) \leqslant \mathsf{t}^+(x - y) \ . \tag{N_{t}^+}$$

Gunawardena and Keane showed in [GK95] that

$$\begin{aligned} \text{(M) and (AH)} &\iff (N_{\mathsf{t}}) \ , \\ \text{(M) and (ASH)} &\iff (N_{\mathsf{t}}^+) \ . \end{aligned}$$

Again, these relations are readily generalized to any vector space $\mathcal{V}$ with an order unit e, defining the function $\mathsf{t} : \mathcal{V} \to \mathbb{R}$ by

$$\mathsf{t}(x) := \inf\{\lambda \in \mathbb{R} \mid x \leqslant \lambda e\} \ , \quad x \in \mathcal{V} \ . \tag{3.2}$$

3.3 Minimax representation theorems

In this section, we give a general minimax characterization of nonexpansive functions with respect to a weak Minkowski norm, which will lead to extension and refinements of the known minimax representations of Shapley operators.

3.3.1 Nonexpansive functions with respect to weak Minkowski norms

We shall need the following lemma, which is a variation of Legendre-Fenchel duality.

Lemma 3.2. *Let $\mathcal{V}$ be a real t.v.s. and let $q : \mathcal{V} \to \mathbb{R}$ be a weak Minkowski norm, continuous with respect to the topology of $\mathcal{V}$. Then, for every vector $x \in \mathcal{V}$,*

$$q(x) = \max_{p \in \mathcal{P}} \langle p, x \rangle \enspace , \tag{3.3}$$

where $\mathcal{P} := \{p \in \mathcal{V}^* \mid \forall x \in \mathcal{V},\ \ \langle p, x \rangle \leqslant q(x)\}$ is a nonempty convex set of $\mathcal{V}'$, compact for the weak- topology.*

Proof. For any $x \in \mathcal{V} \setminus \{0\}$ we have, by definition of $\mathcal{P}$, $q(x) \geqslant \sup_{p \in \mathcal{P}} \langle p, x \rangle$. Let p be the linear form defined on the vector subspace of $\mathcal{V}$ generated by x, $\mathbb{R}x$, and such that $\langle p, x \rangle = q(x)$. If $\lambda \geqslant 0$, then

$$\langle p, \lambda x \rangle = \lambda \langle p, x \rangle = \lambda q(x) = q(\lambda x) \enspace ,$$

since q is positively homogeneous. If $\lambda < 0$, then

$$\langle p, \lambda x \rangle = \lambda \langle p, x \rangle = -|\lambda| q(x) \leqslant |\lambda| q(-x) = q(\lambda x) \enspace ,$$

the inequality coming from the subadditivity of the weak norm: $0 = q(0) \leqslant q(x) + q(-x)$. Hence p is dominated by q on the vector subspace $\mathbb{R}x$, and according to the Hahn-Banach extension theorem [AB06, Th. 5.53], there is a linear extension $\hat{p}$ of p to $\mathcal{V}$ which is dominated by q on $\mathcal{V}$. Therefore, $\hat{p} \in \mathcal{P}$ and $q(x) = \langle \hat{p}, x \rangle$, which shows that $q(x) = \max_{p \in \mathcal{P}} \langle p, x \rangle$. This also proves that $\mathcal{P}$ is nonempty.

Note that, since q is continuous, every linear form p dominated by q is also continuous, hence $\mathcal{P} \subset \mathcal{V}'$. The convexity of $\mathcal{P}$ is straightforward and since this set can also be written as

$$\mathcal{P} = \bigcap_{x \in \mathcal{V}} \{p \in \mathcal{V}^* \mid \langle p, x \rangle \leqslant q(x)\} \enspace ,$$

we deduce that $\mathcal{P}$ is a weak-* closed set. Furthermore, for every $x \in \mathcal{V}$ and every $p \in \mathcal{P}$ we have $-q(-x) \leqslant \langle p, x \rangle \leqslant q(x)$. Hence, $\mathcal{P}$ is pointwise bounded. Applying the Tychonoff Product theorem [AB06, Th. 2.61], we deduce that $\mathcal{P}$ is weak-* compact. □

Remark 3.3. In Lemma 3.2, we did not assume that $\mathcal{V}$ is locally convex or Hausdorff. When these properties hold, Lemma 3.2 becomes a direct application of [AB06, Th. 7.52], which applies more generally to dual pairs. Alternatively, this result can be obtained using the Legendre-Fenchel duality for convex proper lower semicontinuous functions [ET99, Prop. 4.1], still assuming that $\mathcal{V}$ is locally convex and Hausdorff.

The following simple observation shows that the maximum in (3.3) is attained in the closure of the set of extreme points of $\mathcal{P}$.

Lemma 3.4. *Let $q : \mathcal{V} \to \mathbb{R}$ be a continuous weak Minkowski norm, and define $\mathcal{P}$ as in Lemma 3.2. Denote by $\operatorname{ext}\mathcal{P}$ the set of extreme points of $\mathcal{P}$, and by $\overline{\operatorname{ext}\mathcal{P}}$ its closure with respect to the weak-* topology. Then, for every $x \in \mathcal{V}$,*

$$q(x) = \sup_{p \in \operatorname{ext}\mathcal{P}} \langle p, x\rangle = \max_{p \in \overline{\operatorname{ext}\mathcal{P}}} \langle p, x\rangle \enspace . \tag{3.4}$$

In particular, if $\operatorname{ext}\mathcal{P}$ is closed, then the maximum in (3.3) is attained at an extreme point of P.

Proof. We already showed in Lemma 3.2 that for every $x \in \mathcal{V}$, the supremum $q(x) = \sup_{p\in\mathcal{P}}\langle p, x\rangle$ is attained at some linear form $r \in \mathcal{P}$. It follows from the Krein-Milman theorem that $\mathcal{P}$ is the closed convex hull of the set $\operatorname{ext}\mathcal{P}$, the closure being understood with respect to the weak-* topology. Hence, there exists a net (r_α) of elements of the convex hull of $\operatorname{ext}\mathcal{P}$ that converges to r in this topology. In particular, every r_α can be written as a finite sum $r_\alpha = \sum_{1\leqslant i\leqslant m} \lambda_i p_i$ with $\sum_{1\leqslant i\leqslant m} \lambda_i = 1$, $\lambda_i \geqslant 0$ and $p_i \in \operatorname{ext}\mathcal{P}$ for all $i \in [m]$ and some integer $m \geqslant 1$, and so, $\langle r_\alpha, x\rangle = \sum_{1\leqslant i\leqslant m} \lambda_i\langle p_i, x\rangle \leqslant \sup_{p\in\operatorname{ext}\mathcal{P}}\langle p, x\rangle$. We deduce that $q(x) = \lim_\alpha\langle r_\alpha, x\rangle \leqslant \sup_{p\in\operatorname{ext}\mathcal{P}}\langle p, x\rangle$. The opposite inequality $q(x) \geqslant \sup_{p\in\operatorname{ext}\mathcal{P}}\langle p, x\rangle$ follows readily from (3.3). Using the weak-* continuity of the map $p \mapsto \langle p, x\rangle$, we get $\sup_{p\in\operatorname{ext}\mathcal{P}}\langle p, x\rangle = \sup_{p\in\overline{\operatorname{ext}\mathcal{P}}}\langle p, x\rangle$. Finally, since $\overline{\operatorname{ext}\mathcal{P}}$ is a weak-* closed subset of the weak-* compact set $\mathcal{P}$, it is also weak-* compact. So, we deduce that the latter supremum is attained. □

We deduce from the previous lemmas a minimax representation theorem that directly extends the result of Rubinov and Singer [RS01b, Th. 5.3].

Theorem 3.5. *Let $\mathcal{V}$ be a real t.v.s. and let $q : \mathcal{V} \to \mathbb{R}$ be a weak Minkowski norm, continuous with respect to the topology of $\mathcal{V}$. Let $\mathcal{P} := \{p \in \mathcal{V}^* \mid \forall x \in \mathcal{V},\ \langle p, x\rangle \leqslant q(x)\}$. A function $f : \mathcal{V} \to \mathbb{R}$ is nonexpansive with respect to q if, and only if,*

$$f(x) = \min_{y\in\mathcal{V}} \max_{p\in\mathcal{P}} \big(\langle p, x - y\rangle + f(y)\big) \enspace , \quad \forall x \in \mathcal{V} \enspace . \tag{3.5}$$

Moreover, the value of the latter expression does not change if the maximum is restricted to the points $p \in \overline{\operatorname{ext}\mathcal{P}}$.

Proof. If f is nonexpansive with respect to q, then by definition we have $f(x) - f(y) \leqslant q(x - y)$ for all $x, y \in \mathcal{V}$. We readily deduce that $f(x) = \min_{y\in\mathcal{V}} q(x - y) + f(y)$, the minimum being attained in x. Replacing q by its expression given in (3.3), we get the minimax representation formula (3.5). The remaining part of the theorem follows directly from Lemma 3.4.

Conversely, as a consequence of Lemma 3.2, any real function given by (3.5) is easily seen to be nonexpansive with respect to q. □

Corollary 3.6. *Let $\mathcal{V}$ be a real vector space with an order unit, denoted by e, and the topology defined by the sup-norm (3.1) associated with that order unit. Let*

$$\Delta := \{p \in \mathcal{V}^* \mid \langle p, e\rangle = 1,\ \langle p, x\rangle \geqslant 0,\ \forall x \in \mathcal{V}_+\} \enspace ,$$

where $\mathcal{V}_+ := \{x \in \mathcal{V} \mid x \geqslant 0\}$. Then, a function $f : \mathcal{V} \to \mathbb{R}$ is monotone and additively homogeneous if, and only if,

$$f(x) = \min_{y\in\mathcal{V}} \max_{p\in\Delta} \big(\langle p, x - y\rangle + f(y)\big) \enspace .$$

Moreover, the value of the latter expression is not changed if the minimum is restricted to the vectors $y \in \mathcal{V}$ such that $f(y) = 0$, or if the maximum is restricted to the linear forms $p \in \overline{\operatorname{ext}\Delta}$.

Proof. Let f be a monotone and additively homogeneous real function on $\mathcal{V}$. As exposed in Section 3.2, this is equivalent to the function f being nonexpansive with respect to the weak Minkowski norm $\mathrm{t}(\cdot)$ defined in (3.2). Let $\mathcal{P}$ be defined as in Theorem 3.5. We next show that $\mathcal{P} = \Delta$.

If $p \in \mathcal{P}$, then for all $x \in \mathcal{V}_+$ we have $\mathrm{t}(-x) \leqslant 0$, so that $\langle p, -x\rangle \leqslant 0$, hence $\langle p, x\rangle \geqslant 0$. Moreover, $\mathrm{t}(e) = 1$ and $\mathrm{t}(-e) = -1$, so that $\langle p, e\rangle \leqslant 1$ and $\langle p, -e\rangle \leqslant -1$, which shows that $\langle p, e\rangle = 1$. This shows that $\mathcal{P} \subset \Delta$. Conversely, if $p \in \Delta$, then for all $x \in \mathcal{V}$ such that $\mathrm{t}(x) < \lambda$, we have $x \leqslant \lambda e$, so that $\langle p, x\rangle = \langle p, x - \lambda e\rangle + \lambda\langle p, e\rangle \leqslant \lambda$. This implies that $\langle p, x\rangle \leqslant \mathrm{t}(x)$ for all $x \in \mathcal{V}$, hence $p \in \mathcal{P}$. Applying Theorem 3.5, we get the first equality in the corollary.

Now, remark that $\langle p, x - y\rangle + f(y) = \langle p, x - z\rangle$ with $z = y - f(y)e$ and $f(z) = 0$. Then, a change of variable leads to

$$f(x) = \min_{\substack{y\in\mathcal{V} \\ f(y)=0}} \max_{p\in\Delta} \langle p, x - y\rangle \enspace .$$

The remaining part of the corollary follows from Theorem 3.5 and the converse is straightforward. □

Example 3.7. If $\mathcal{V} = \mathbb{R}^n$ and $F : \mathcal{V} \to \mathcal{V}$ is monotone and additively homogeneous, then, as recalled in Section 3.2, each coordinate function F_i, $i \in [n]$, is nonexpansive with respect to the weak Minkowski norm $\mathrm{t}(\cdot)$. Then, the representation result of Rubinov and Singer [RS01b, Th. 5.3], which shows that

$$F_i(x) = \min_{y\in\mathbb{R}^n} \max_{j\in[n]} \big(x_j - y_j + F_i(y)\big) \enspace , \quad \forall x \in \mathbb{R}^n \enspace ,$$

is a special case of Corollary 3.6.

3.3.2 Positively homogeneous nonexpansive functions

We now consider nonexpansive functions that are positively homogeneous. The following theorem characterizes these functions as minimax of nonexpansive linear forms.

Theorem 3.8. *Let $\mathcal{V}$ be a real t.v.s. and let $q : \mathcal{V} \to \mathbb{R}$ be a weak Minkowski norm, continuous with respect to the topology of $\mathcal{V}$. Denote by $\mathcal{P} := \{p \in \mathcal{V}^* \mid \forall x \in \mathcal{V},\ \langle p, x\rangle \leqslant q(x)\}$. A function $f : \mathcal{V} \to \mathbb{R}$ is positively homogeneous and nonexpansive with respect to q if, and only if,*

$$f(x) = \min_{y\in\mathcal{V}} \max_{p\in\mathcal{P}_y} \langle p, x\rangle \enspace , \tag{3.6}$$

where $\mathcal{P}_y := \{p \in \mathcal{P} \mid \langle p, y\rangle \leqslant f(y)\}$. Moreover, the value of the expression (3.6) does not change if the maximum is restricted to the linear forms $p \in \overline{\operatorname{ext} \mathcal{P}_y}$.

Proof. The "if" part of the condition is straightforward, so we only prove that it is necessary. Let f be a real function positively homogeneous and nonexpansive with respect to q. As a direct consequence of nonexpansiveness, we get that

$$f(x) = \min_{y\in\mathcal{V}} \big(f(y) + q(x - y)\big) \enspace .$$

Since f is also positively homogeneous, by a change of variable, we have

$$f(x) = \min_{y\in\mathcal{V}} \inf_{\lambda>0} \big(\lambda f(y) + q(x - \lambda y)\big) \enspace .$$

There, the minimum in y is attained at all μx with $\mu > 0$. Given a vector $y \in \mathcal{V}$, let g_y be the real function defined on $\mathcal{V}$ by $g_y(x) = \inf_{\lambda>0} \lambda f(y) + q(x - \lambda y)$, so that $f(x) = \min_{y\in\mathcal{V}} g_y(x)$. We next show that g_y is a continuous weak Minkowski norm.

First, g_y is bounded above by q and below by f. In particular, g_y is bounded above on a neighborhood of each point of $\mathcal{V}$. Second, using the positive homogeneity of f and q, we check that g_y also satisfies this property. Third, g_y is convex. Indeed, the function $x \mapsto f(y) + q(x - y)$ is convex because so is q. Its *perspective function*, defined on $\mathbb{R} \times \mathcal{V}$ by $(\lambda, x) \mapsto \lambda f(y) + q(x - \lambda y)$ if $\lambda > 0$ and $+\infty$ otherwise, is also convex, see [BC11, Prop. 8.23]. Hence, the convexity follows from the fact that g_y is the *marginal function* with respect to the first variable of the latter function.

We have shown that g_y is a positively homogeneous convex function, finite everywhere and bounded above on a neighborhood of each point of $\mathcal{V}$. Hence, it is continuous on $\mathcal{V}$, see [AB06, Th. 5.43], and we deduce from Lemma 3.2 that it is the *support function* of the weak-* compact convex set $\mathcal{Q}_y := \{p \in \mathcal{V}^* \mid \forall x \in \mathcal{V},\ \langle p, x\rangle \leqslant g_y(x)\}$. To conclude, it remains to show that this set is $\mathcal{P}_y$.

Let $p \in \mathcal{Q}_y$. Then, for every $x \in \mathcal{V}$ and every $\lambda > 0$, we have $\langle p, x\rangle \leqslant \lambda f(y) + q(x - \lambda y)$. Taking $\lambda \to 0$ we deduce that $p \in \mathcal{P}$, and taking $x = y$ and $\lambda = 1$ we deduce that $\langle p, y\rangle \leqslant f(y)$. Hence, $p \in \mathcal{P}_y$ which shows that $\mathcal{Q}_y \subset \mathcal{P}_y$. Consider now $p \in \mathcal{P}_y$. For every $x \in \mathcal{V}$ and every $\lambda > 0$ we have

$$\langle p, x\rangle = \langle p, x\rangle - \langle p, \lambda y\rangle + \langle p, \lambda y\rangle = \langle p, x - \lambda y\rangle + \lambda\langle p, y\rangle \leqslant q(x - \lambda y) + \lambda f(y) \ .$$

Taking the infimum over all $\lambda > 0$ in the right-hand side of the last inequality, we get that $\langle p, x\rangle \leqslant g_y(x)$. Hence $p \in \mathcal{Q}_y$ which shows that $\mathcal{Q}_y \subset \mathcal{P}_y$ and consequently that $\mathcal{Q}_y = \mathcal{P}_y$. □

The following is an immediate corollary.

Corollary 3.9. *Let $\mathcal{V}$ be a real vector space with an order unit, denoted by e, and the topology defined by the sup-norm (3.1) associated with that order unit. Then, a function $f : \mathcal{V} \to \mathbb{R}$ is monotone, additively homogeneous, and positively homogeneous if, and only if,*

$$f(x) = \min_{y \in \mathcal{V}} \max_{p \in \Delta_y} \langle p, x\rangle \ , \tag{3.7}$$

where

$$\Delta_y := \{p \in \mathcal{V}^* \mid \langle p, y\rangle \leqslant f(y),\ \ \langle p, e\rangle = 1,\ \ \langle p, x\rangle \geqslant 0,\ \ \forall x \in \mathcal{V}_+\} \ .$$

Moreover, the value of the expression in (3.7) does not change if the minimum is restricted to the vectors $y \in \mathcal{V}$ such that $f(y) = 0$, or if the maximum is restricted to the linear forms $p \in \overline{\operatorname{ext} \Delta_y}$.

3.4 Applications

We now point out some applications of the present representation theorems to nonconvex risk measures and Shapley operators.

3.4.1 Representation of nonconvex risk measures

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. Denote by $L^\infty(\mathbb{P})$ the space of almost surely bounded real-valued random variables, equipped with the usual essential sup-norm. This is a special case of an AM-space with unit, the unit $\mathbb{1}$ being the random variable a.s. equal to 1. Its topological dual space is the space of finitely additive measures of bounded variation which are absolutely continuous with respect to $\mathbb{P}$, see [DS88, Th. IV.8.16]. We denote it by $\mathsf{ba}(\mathbb{P})$. Following standard notation, we shall write $\mathbb{E}_p[X]$ instead of $\langle p, X\rangle$ for a random variable $X \in L^\infty(\mathbb{P})$ and a measure $p \in \mathsf{ba}(\mathbb{P})$. We denote by $\mathsf{ba}^+(\mathbb{P})$ the set of positive bounded finitely additive measures and by $\Delta(\mathbb{P}) := \{p \in \mathsf{ba}^+(\mathbb{P}) \mid \mathbb{E}_p[\mathbb{1}] = 1\}$ the set of finitely additive probability measures.

A *risk measure* is a real function $\mu : L^\infty(\mathbb{P}) \to \mathbb{R}$ that satisfies the following conditions, for all $X, Y \in L^\infty(\mathbb{P})$:

(i) $X \leqslant Y \implies \mu(X) \geqslant \mu(Y)$;
(ii) $\mu(X + \lambda) = \mu(X) - \lambda, \quad \forall \lambda \in \mathbb{R}$;
see [FS11]. Observe that this is equivalent to the function $\rho : X \mapsto \mu(-X)$ being monotone and additively homogeneous.

A risk measure μ is *coherent* if it is in addition convex and positively homogeneous, see e.g. [ADEH99, Del02]. It is known that a coherent risk measure can be represented in the form

$$\mu(X) = \sup_{p \in \mathcal{Q}_{\mathsf{ba}}} \mathbb{E}_p[-X]$$

where $\mathcal{Q}_{\mathsf{ba}} \subset \Delta(\mathbb{P})$ is a convex subset of the space of finitely additive probability measures on Ω, closed with respect to the $\sigma(\mathsf{ba}(\mathbb{P}), L^\infty(\mathbb{P}))$-topology. As a direct application of Corollary 3.6, we obtain a similar representation for general nonconvex risk measures.

Corollary 3.10. *Let μ be a risk measure on $L^\infty(\mathbb{P})$. Then,*

$$\mu(X) = \min_{\substack{Y \in L^\infty(\mathbb{P}) \\ \mu(Y)=0}} \max_{p \in \Delta(\mathbb{P})} \mathbb{E}_p[Y - X] \enspace .$$

The following corollary characterizes the nonconvex risk measures that are positively homogeneous. It is a direct application of Corollary 3.9.

Corollary 3.11. *Let μ be a positively homogeneous risk measure. Then,*

$$\mu(X) = \min_{\substack{Y \in L^\infty(\mathbb{P}) \\ \mu(Y)=0}} \max_{\substack{p \in \Delta(\mathbb{P}) \\ \mathbb{E}_p[Y] \geqslant 0}} \mathbb{E}_p[-X] \enspace .$$

3.4.2 Representation of payment-free Shapley operators

Here, we consider the vector space $\mathcal{V} = \mathbb{R}^n$, where $\langle \cdot, \cdot \rangle$ denotes the standard scalar product. Recall that a *payment-free Shapley operator* on $\mathbb{R}^n$ is a monotone, additively and positively homogeneous operator from $\mathbb{R}^n$ to itself. As mentioned in the introduction, this terminology is justified by the following corollary, which shows that all operators of this kind precisely arise from zero-sum stochastic games the payment function of which is identically zero.

Corollary 3.12. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a monotone, additively and positively homogeneous operator. Then, each coordinate map F_i with $i \in [n]$ can be represented as*

$$F_i(x) = \min_{a \in A_i} \max_{b \in B_{i,a}} \sum_{j \in [n]} x_j \, p(j \mid i, a, b) \enspace , \quad \forall x \in \mathbb{R}^n \enspace , \tag{3.8}$$

where $A_i := \{a \in \mathbb{R}^n \mid F_i(a) = 0\}$; for every $i \in [n]$ and every $a \in A_i$, $B_{i,a}$ is a finite subset of $\mathbb{R}^n$; for every $a \in A_i$ and every $b \in B_{i,a}$, $(p(j \mid i, a, b))_{j \in [n]}$ is a stochastic vector with at most two positive coordinates.

Proof. Each coordinate function of F is monotone, additively and positively homogeneous. Then, it follows from Corollary 3.9 that, for all $i \in [n]$ and all $x \in \mathbb{R}^n$,

$$F_i(x) = \min_{\substack{y \in \mathbb{R}^n \\ F_i(y)=0}} \max_{\substack{p \in \Delta_n \\ \langle p, y \rangle \leqslant 0}} \langle p, x \rangle \enspace . \tag{3.9}$$

where Δ_n is the standard simplex of $\mathbb{R}^n$.

Let $A_i := \{a \in \mathbb{R}^n \mid F_i(a) = 0\}$. For $a \in A_i$, let $B_{i,a}$ be the set of extreme points of the polytope $\{p \in \Delta_n \mid \langle p, a \rangle \leqslant 0\}$ (points at which the maximum in (3.9) is attained). Finally, for

$a \in A_i$, $b \in B_{i,a}$ and $j \in [n]$, let $p(j \mid i, a, b) = b$. Rewriting equation (3.9) using these notations we get exactly the formula (3.8).

A standard result of convex geometry shows that every extreme point of the intersection of a polytope with a half-space is either an extreme point of the polytope, or a convex combination of two extreme points of this polytope (see for instance [FP96, Lem. 3]). It follows that every element in $B_{i,a}$ is either an extreme point of Δ_n or a convex combination of two extreme points of Δ_n. □

3.4.3 Approximation of Shapley operators

We use the previous result to approximate payment-free Shapley operators on $\mathbb{R}^n$ by minimax maps where the min and max operators are taken over finite sets (i.e., continuous piecewise affine operators). The latter maps play an important role algorithmically, in max-plus finite element method [AGL08], and more generally in idempotent methods [McE11, MP15].

Let $q : \mathbb{R}^n \to \mathbb{R}$ be a weak Minkowski norm. We say that a subset $A \subset \mathbb{R}^n$ is an ε-net of any set $K \subset \mathbb{R}^n$ with respect to (the symmetrization of) q if

$$\inf_{a \in A} \max\{q(x-a), q(a-x)\} < \varepsilon \ , \quad \forall x \in K \ .$$

Note that here, since the dimension is finite, q is continuous, and then it is always possible to find a finite ε-net of any compact set with respect to q.

The following result, when applied to any Shapley operators on $\mathbb{R}^n$, shows that the latter can be approximated on any compact set by the Shapley operators of a repeated game with deterministic transitions and finite action spaces.

Proposition 3.13. *Let $q : \mathbb{R}^n \to \mathbb{R}$ be a weak Minkowski norm and let $f : \mathbb{R}^n \to \mathbb{R}$ be a map nonexpansive with respect to q. Then, for every subset $K \subset \mathbb{R}^n$, and for every finite ε-net $(y_\ell)_{\ell \in [m]}$ of K with respect to q, the function*

$$g(x) = \min_{\ell \in [m]} \big(f(y_\ell) + q(x - y_\ell) \big)$$

is such that

$$f(x) \leqslant g(x) \leqslant f(x) + 2\varepsilon \ , \quad \forall x \in K \ .$$

Proof. Let $x \in K$. Since f is nonexpansive with respect to q, we have $f(x) \leqslant f(y_\ell) + q(x - y_\ell)$ for all indices $\ell \in [m]$. Thus, we have $f(x) \leqslant g(x)$. We also know that there exists an index ℓ_0 such that $\max\{q(x - y_{\ell_0}), q(y_{\ell_0} - x)\} < \varepsilon$. Hence,

$$f(y_{\ell_0}) + q(x - y_{\ell_0}) \leqslant f(x) + q(y_{\ell_0} - x) + q(x - y_{\ell_0}) \leqslant f(x) + 2\varepsilon \ ,$$

from which the second inequality follows. □

Note that in the previous proposition, the polyhedral function g is not positively homogeneous, even if f is. The next result shows that when a Shapley operator is payment-free, it can be approximated by another payment-free Shapley operator.

Corollary 3.14. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a payment-free Shapley operator. Then, for all $\varepsilon > 0$, there exists a payment-free Shapley operator $G : \mathbb{R}^n \to \mathbb{R}^n$ such that*

$$F(x) \leqslant G(x) \leqslant F(x) + \varepsilon \|x\|_\infty \, e \ , \quad \forall x \in \mathbb{R}^n \ ,$$

which can be represented in the form

$$G_i(x) = \min_{a \in A_i} \max_{b \in B_{i,a}} \sum_{j \in [n]} x_j \, p(j \mid i, a, b) \quad , \quad \forall x \in \mathbb{R}^n, \quad \forall i \in [n] \quad ,$$

where A_i and $B_{i,a}$ are finite *sets, and every $(p(j \mid i, a, b))_{j \in [n]}$ is a stochastic vector with at most two positive coordinates.*

Proof. Let $(z_\ell)_{\ell \in [m]}$ be an $\varepsilon/2$-net of the unit sphere of $\mathbb{R}^n$ with respect to the sup-norm. Applying Corollary 3.12 to F, we know that for all $i \in [n]$ and $\ell \in [m]$, there exists an action $a_{i\ell} \in A_i$ for which the minimum is attained in the minimax representation (3.8) of $F_i(z_\ell)$. Let $A_i^* := \{a_{i\ell} \mid \ell \in [m]\}$ be the finite subset of A_i containing all these optimal actions in state $i \in [n]$. Then, let $G : \mathbb{R}^n \to \mathbb{R}^n$ be the payment-free Shapley operator the ith coordinate map of which is given by

$$G_i(x) := \min_{a \in A_i^*} \max_{b \in B_{i,a}} \sum_{j \in [n]} x_j \, p(j \mid i, a, b) \quad , \quad \forall x \in \mathbb{R}^n \quad ,$$

where the action spaces $B_{i,a}$ are the same than in the minimax representation (3.8) of F_i. In particular, we know that they are finite and that all the vectors $(p(j \mid i, a, b))_{j \in [n]}$ are stochastic, with at most two positive coordinates.

By construction, we have $F(x) \leqslant G(x)$ for all vectors $x \in \mathbb{R}^n$, with equality for every z_ℓ, $\ell \in [m]$. Now, given a vector x in the unit sphere of $\mathbb{R}^n$, we can choose a vector z_ℓ such that $\|x - z_\ell\|_\infty \leqslant \varepsilon/2$. Since both F and G are nonexpansive with respect to the sup-norm, we deduce that for all $i \in [n]$,

$$G_i(x) - F_i(x) \leqslant G_i(z_\ell) + \varepsilon/2 - F_i(x) = F_i(z_\ell) - F_i(x) + \varepsilon/2 \leqslant \varepsilon \quad .$$

The conclusion follows from the positive homogeneity of F and G. □

Part II

Ergodicity of zero-sum stochastic games

CHAPTER 4

Analytic characterizations of ergodicity of stochastic games

4.1 Introduction

The problem of existence of the mean-payoff vector for a stochastic game with finite state space may be addressed by finding a solution to the ergodic equation (2.5), involving its Shapley operator T. The latter equation has a solution if, and only if, some orbit of T is bounded in Hilbert's seminorm (Theorem 2.6). Finding conditions under which this property is true is difficult in general, and depends not only on the structure of the game (i.e., the support of the transition probabilities), but also on the values of the parameters (payment and transition functions). Following Gaubert and Gunawardena [GG04], a better posed problem consists in finding conditions under which all the subsets of a given family of invariant spaces are bounded in Hilbert's seminorm. Such conditions usually depend only on the structure of the game (see Theorem 2.8 for instance). Then the solvability of the ergodic equation is ensured not only for T, but also for some family of perturbations of its parameters.

An elementary case illustrating this stability property is given by ergodic finite Markov chains. Following Kemeny and Snell [KS76], we define the ergodicity of a finite Markov chain by one of the equivalent properties listed in the theorem below, involving its transition matrix. Recall in particular that for a $n \times n$ stochastic matrix P, the *directed graph* associated with P is composed of the nodes $1, \dots, n$ and of the arcs (i, j) such that $P_{ij} > 0$. A *class* of the matrix P is a maximal set of nodes such that every two nodes in this set are connected by a directed path. A class is said to be *final* if every path starting from a node of this class remains in it. We refer the reader to [KS76] and [BP94] for the proof of the theorem.

Theorem 4.1. *Let $P \in \mathbb{R}^{n \times n}$ be a stochastic matrix. The following assertions are equivalent:*

(i) every vector $v \in \mathbb{R}^n$ such that $Pv = v$ is constant;
(ii) for all vectors $g \in \mathbb{R}^n$, there exists an eigenpair $(\lambda, u) \in \mathbb{R} \times \mathbb{R}^n$ such that $g + Pu = \lambda e + u$;
(iii) for all vectors $g \in \mathbb{R}^n$, the Cesàro limit $\lim_{k \to \infty}(g + Pg + \dots + P^{k-1}g)/k$ is a constant vector;
(iv) the directed graph associated with the matrix P has a unique final class;
(v) the stochastic matrix P has a unique invariant probability measure.

A finite Markov chain with transition matrix $P \in \mathbb{R}^{n \times n}$ may be seen as a zero-player stochastic game with state space $[n]$. If g_i is the payment in state $i \in [n]$, then its Shapley operator is given by $T(x) = g + Px$ for all $x \in \mathbb{R}^n$, and it is readily seen that its recession operator is equal to $\widehat{T}(x) = Px$. Thus, the ergodicity of a finite Markov chain may be characterized, in game-theoretic terms, either by all the fixed points of the recession operator $\widehat{T}$ being constant (Point (i)), or by the ergodic equation having a solution for all Shapley operators $g + T$ with $g \in \mathbb{R}^n$ (Point (ii)), or by the mean-payoff vector being constant for all payment vectors $g \in \mathbb{R}^n$ (Point (iii)). A natural question is to know if this result may be extended to any finite stochastic game.

Let us recall that the stability of existence of a constant mean payoff, based on structural properties, was observed and studied in various framework. In the case of finite Markov chains with risk-sensitive long-run average cost criterion, Cavazos-Cadena and Hernández-Hernández [CCHH09] gave a necessary and sufficient condition on the transition matrix for the Poisson equation – a dynamic programming type equation – to have a solution for any cost function. Still motivated by risk-sensitive applications, the same authors [CCHH10] gave, under a weakly convex property and in a multiplicative framework, necessary and sufficient conditions for the nonlinear spectral problem (2.6) to have a solution for arbitrary (multiplicative) perturbations of the payments. Also, in optimal control (with continuous time and infinite state space) Arisawa [Ari97, Ari98] studied the ergodic problem of Hamilton-Jacobi-Bellman equations and particularly the link between the existence of the ergodic constant for arbitrary cost functions depending only on the state and the existence of an "ergodic attractor", depending on the dynamic.

In this chapter and the next one, we extend the notion of ergodicity to zero-sum stochastic games. This leads us to the following definition.

Definition 4.2. A Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is *ergodic* if the ergodic equation (2.5) is solvable for all operators $g + T$ with $g \in \mathbb{R}^n$. A zero-sum stochastic game with finite state space is *ergodic* if its Shapley operator is ergodic.

In this chapter, along the lines of Gaubert and Gunawardena [GG04], we start by characterizing ergodicity in terms of recession operator, assuming that the asymptotic behavior of the Shapley operator and its recession operator are the same. Then we characterize ergodicity in full generality, in terms of slice spaces, using the connections between nonexpansive and accretive maps. In the following two chapters we deal with the combinatorial aspects of these conditions. In particular, we formulate them in terms of hypergraphs, underlying their structural aspects.

The notion of ergodicity for stochastic games and the results of Section 4.2 are based on the first part of the paper [AGH14a], published in *Discrete and Continuous Dynamical Systems, Series A*. The results regarding the accretive operator approach (Section 4.3 and 4.4) have been announced in the MTNS conference proceedings [Hoc16].

4.2 Characterization of ergodicity in terms of recession operator

Let T be a Shapley operator on $\mathbb{R}^n$. Assume that its recession operator $\widehat{T}$ exists. We know that if all the fixed points of $\widehat{T}$ are trivial (i.e., constant vectors), then all the slice spaces of T are bounded in Hilbert's seminorm (Theorem 2.9). This implies that the ergodic equation (2.5) is solvable and that the mean-payoff vector is constant. Note that it remains true for all operators $g + T$ with $g \in \mathbb{R}^n$, since they all have the same recession operator. Hence, T is ergodic.

The following proposition specifies the link between the recession operator and the (not necessarily constant) mean-payoff vector, when they do exist. We give the short proof for convenience of the reader.

Proposition 4.3 ([RS01a]). *Let T be a Shapley operator on $\mathbb{R}^n$. Assume that its recession operator $\widehat{T}$ and the mean-payoff vector $\chi \in \mathbb{R}^n$ exist. Then, $\widehat{T}(\chi) = \chi$.*

Proof. Since T is nonexpansive in the sup-norm, we have, for all vectors $x, y \in \mathbb{R}^n$ and all integer k,

$$\left\| \frac{T(kx) - T(ky)}{k} \right\| \leqslant \|x - y\| \ .$$

Then, letting $x = \chi$ and $y = T^k(0)/k$, we get

$$\left\| \frac{T(k\chi)}{k} - \frac{T^{k+1}(0)}{k} \right\| \leqslant \left\| \chi - \frac{T^k(0)}{k} \right\| \ .$$

Now, taking the limits in k, we obtain that $\|\widehat{T}(\chi) - \chi\| = 0$. □

A converse statement holds.

Proposition 4.4. *Let F be a payment-free Shapley operator on $\mathbb{R}^n$, and let v be a fixed point of F. Then, there is a Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ such that its recession operator exists and is given by $\widehat{T} = F$, and such that its mean-payoff vector is equal to v.*

Proof. Let $T := v + F$. It is readily seen that $\widehat{T} = F$. Moreover, for every integer k, we have $T(kv) = (k+1)v$, so that, by induction, $T^k(0) = kv$. Hence the mean-payoff vector associated with T exists and is equal to v. □

Remark 4.5. Let $\Gamma(r) = ([n], A, B, K_A, K_B, r, p)$ be a parametric stochastic game whose parameter r is chosen among all bounded real functions on K. It is readily seen that all Shapley operators of the games $\Gamma(r)$ have the same recession operator, denoted by F. The latter is the (payment-free) Shapley operator of the game $\Gamma(0)$. Then, the fixed points of F give exactly all the realizable mean-payoff vectors of the family of games $(\Gamma(r))$.

We now state the main result of this section, which gives a characterization of ergodicity of a Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ in the sense of Definition 4.2. This characterization relies on the existence of its recession operator. More precisely, we shall need the following assumption.

Assumption 4.A. The recession operator $\widehat{T} : \mathbb{R}^n \to \mathbb{R}^n$ is well defined and there exists a real constant $M > 0$ such that

$$\left| T_i(x) - \widehat{T}_i(x) \right| \leqslant M \ , \quad \forall x \in \mathbb{R}^n \ , \ \ \forall i \in [n] \ .$$

Note that Assumption 4.A is satisfied as soon as T is the Shapley operator of a stochastic game with bounded payment function, the constant M being a bound on the payments. Indeed, if $T : \mathbb{R}^n \to \mathbb{R}^n$ is given by formula (2.1), then the ith component of its recession operator is readily seen to be

$$\widehat{T}_i(x) = \inf_{\mu \in \Delta(A_i)} \sup_{\nu \in \Delta(B_i)} \sum_{j \in [n]} x_j \, p(j \mid i, \mu, \nu) \ , \quad x \in \mathbb{R}^n \ ,$$

where the infimum and the supremum commutes. Moreover, Assumption 4.A implies that the convergence in the limit (2.7) defining $\widehat{T}$ is uniform in x.

Theorem 4.6. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a Shapley operator satisfying Assumption 4.A. The following assertion are equivalent:*

(i) $\widehat{T}$ has only trivial fixed points;

(ii) T is ergodic, i.e., for all vectors $g \in \mathbb{R}^n$, the ergodic equation (2.5) is solvable for $g + T$;
(iii) for all vectors $g \in \mathbb{R}^n$, the mean-payoff vector associated with $g + T$ exists and is constant.

Proof. From Theorem 2.9, we know that (i) $\Rightarrow$ (ii) and the implication (ii) $\Rightarrow$ (iii) is straightforward. Hence, all the equivalences will follow from the implication (iii) $\Rightarrow$ (i).

Assume that Point (iii) holds and let v be a fixed point of $\widehat{T}$. Let ℓ be any integer and let $T_\ell := \ell v + T$. We next show by induction that, for all integers k, we have

$$k(\ell v - Me) \leqslant (T_\ell)^k(0) \leqslant k(\ell v + Me) \ . \tag{4.1}$$

For $k = 0$, the inequalities are trivial. Assume now that (4.1) holds for some integer k. Then, by monotonicity of T we have

$$\ell v + T\big(k(\ell v - Me)\big) \leqslant (T_\ell)^{k+1}(0) \leqslant \ell v + T\big(k(\ell v + Me)\big) \ .$$

By additive homogeneity of T we deduce that

$$\ell v - kMe + T(k\ell v) \leqslant (T_\ell)^{k+1}(0) \leqslant \ell v + kMe + T(k\ell v) \ .$$

Using Assumption 4.A, we get that

$$\ell v - (k+1)Me + \widehat{T}(k\ell v) \leqslant (T_\ell)^{k+1}(0) \leqslant \ell v + (k+1)Me + \widehat{T}(k\ell v) \ ,$$

and since v is a fixed point of $\widehat{T}$, the positive homogeneity yields

$$(k+1)(\ell v - Me) \leqslant (T_\ell)^{k+1}(0) \leqslant (k+1)(\ell v + Me) \ .$$

According to Point (iii), the mean-payoff vector of T_ℓ exists and is constant. Let λ_ℓ be this constant value. Then we have $\ell v - Me \leqslant \lambda_\ell e \leqslant \ell v + Me$, which may be rewritten as $(\lambda_\ell - M)e \leqslant \ell v \leqslant (\lambda_\ell + M)e$. This entails $\ell \|v\|_{\mathrm{H}} = \|\ell v\|_{\mathrm{H}} \leqslant 2M$. Since the latter inequality holds for any integer ℓ, we deduce that $\|v\|_{\mathrm{H}} = 0$, which means that v is proportional to the unit vector. □

Note that the above theorem readily extends the case of finite Markov chains, stated in Theorem 4.1 (Points (i) – (iii)). The graph-theoretic aspect of Point (iv) will be treated in the next chapter.

4.3 Accretive operator approach – surjectivity conditions for accretive maps

The purpose of this section and the next one is to extend the characterization of ergodicity to any Shapley operators on $\mathbb{R}^n$. Indeed the conclusions of Theorem 4.6 are not valid in general. For instance, one cannot conclude when the recession operator of T does not exist. But there are also situations where $\widehat{T}$, although well defined, fails to have only trivial fixed points, whereas all slice spaces $\mathcal{S}_\alpha^\beta(T)$ are bounded in Hilbert's seminorm, implying ergodicity. Such a counterexample is provided in Subsection 6.3, others may be found in [GG04].

In the present section, we establish results about the surjectivity of accretive maps in finite dimension. Then, in the next section, we exploit the link between nonexpansiveness and accretivity to infer stability results about the existence of fixed points for nonexpansive maps, including Shapley operators.

4.3.1 Preliminaries

In the remainder of the chapter, $\mathcal{X}$ shall always be a finite-dimensional real vector space. We denote by $\mathcal{X}^*$ its dual space, and by $\langle \cdot, \cdot \rangle$ the duality product. We assume that $\mathcal{X}$ is equipped with a given norm, denoted by $\| \cdot \|$. The dual norm on $\mathcal{X}^*$ is denoted by $\| \cdot \|_*$.

Set-valued maps

Given a space $\mathcal{Y}$, we denote by $A : \mathcal{X} \rightrightarrows \mathcal{Y}$ a *set-valued map* A from $\mathcal{X}$ to $\mathcal{Y}$, i.e., a map from $\mathcal{X}$ to the powerset of $\mathcal{Y}$. We shall need the following definitions.

- The *domain* of A is the subset of $\mathcal{X}$ defined by $\mathrm{dom}(A) := \{x \in \mathcal{X} \mid A(x) \neq \emptyset\}$;
- The *range* of A is the subset of $\mathcal{Y}$ defined by $\mathrm{rg}(A) := \bigcup_{x \in \mathcal{X}} A(x)$;
- The *image* of a subset $\mathcal{V} \subset \mathcal{X}$ under A is the subset of $\mathcal{Y}$ given by $A(\mathcal{V}) := \bigcup_{x \in \mathcal{V}} A(x)$.

The *inverse* of A, denoted by A^{-1}, is the set-valued map from $\mathcal{Y}$ to $\mathcal{X}$ sending an element $y \in \mathcal{Y}$ to the subset $\{x \in \mathcal{X} \mid y \in A(x)\}$, so that $x \in A^{-1}(y)$ if, and only if, $y \in A(x)$. In particular, note that $(A^{-1})^{-1} = A$, and that $\mathrm{rg}(A) = \mathrm{dom}(A^{-1})$. Furthermore, the image of a subset $\mathcal{W} \subset \mathcal{Y}$ under A^{-1} can be written as $A^{-1}(\mathcal{W}) = \{x \in \mathcal{X} \mid A(x) \cap \mathcal{W} \neq \emptyset\}$.

We shall also need the notions of upper and lower semicontinuity. The set-valued map $A : \mathcal{X} \rightrightarrows \mathcal{Y}$ is

- *upper semicontinuous* at point $x \in \mathrm{dom}(A)$ if for every neighborhood $\mathcal{V}$ of $A(x)$, there exists a neighborhood $\mathcal{U}$ of x such that $A(y) \subset \mathcal{V}$ for all $y \in \mathcal{U}$;
- *lower semicontinous* at point $x \in \mathrm{dom}(A)$ if for every open set $\mathcal{V}$ that meets $A(x)$, meaning that $\mathcal{V} \cap A(x) \neq \emptyset$, there exists a neighborhood $\mathcal{U}$ of x such that $A(y) \cap \mathcal{V} \neq \emptyset$ for every $y \in \mathcal{U}$;
- *continuous* at $x \in \mathrm{dom}(A)$ if it is both upper and lower semicontinuous at point x.

The map A is upper (resp., lower) semicontinuous, if it is upper (resp., lower) semicontinuous at every point $x \in \mathrm{dom}(A)$. Likewise, it is *continuous* if it is continuous at every point x in $\mathrm{dom}(A)$. Assume now that $\mathcal{X}$ and $\mathcal{Y}$ are metric spaces. The latter notions may be described in terms of sequences. Let x be a point in the domain of A. The map A is

- upper semicontinous at x and $A(x)$ is compact if, and only if, for every sequence of points $x_k \in \mathrm{dom}(A)$ converging to x and every $y_k \in A(x_k)$, the sequence (y_k) has an accumulation point in $A(x)$;
- lower semicontinuous at x if, and only if, for every $y \in A(x)$ and for every sequence of points $x_k \in \mathrm{dom}(A)$ converging to x, there exists a sequence of points $y_k \in A(x_k)$ converging to y.

We refer the reader to the monograph of Aubin and Frankowska [AF09] for background on set-valued maps.

Normalized duality mapping

The concept of *duality mappings* was introduced by Beurling and Livingston [BL62], and developed later on by Asplund [Asp67], Browder [Bro65, Bro66, Bro76], Browder and de Figueiredo [BdF66]. In a Banach space, there is a strong link between the geometric properties of the space (strict convexity, smoothness, etc.) and the analytic properties of the duality mappings (single-valued, continuity, etc.). Here, we only consider the *normalized duality mapping*, and refer the reader to the above-mentioned papers for a more general presentation.

Definition 4.7. The (normalized) duality mapping on $(\mathcal{X}, \| \cdot \|)$ is the set-valued map $J : \mathcal{X} \rightrightarrows \mathcal{X}^*$ defined by

$$J(x) := \{x^* \in \mathcal{X}^* \mid \|x^*\|_* = \|x\| \ , \ \langle x, x^* \rangle = \|x^*\|_* \|x\|\} \ , \quad x \in \mathcal{X} \ .$$

Note that, by the Hahn-Banach theorem, $\operatorname{dom}(J) = \mathcal{X}$, i.e., $J(x)$ is nonempty for every vector $x \in \mathcal{X}$. Furthermore, Asplund [Asp67, Th. 1] characterized duality mappings as subdifferentials of convex functions. This entails that $J(x)$ is a compact convex subset of $\mathcal{X}^*$ for every $x \in \mathcal{X}$. Also, it readily stems from the definition that J is *homogeneous* of degree one, i.e., for every $x \in \mathcal{X}$ and every $\lambda \in \mathbb{R}$, we have $J(\lambda x) = \lambda J(x)$.

Finally, we shall need the following result, obtained by Browder (see the proof of Th. 7.3 in the reference).

Proposition 4.8 ([Bro76]). *The normalized duality mapping on a finite-dimensional vector space is upper semicontinuous.*

Example 4.9. Let $\mathcal{X} = \mathbb{R}^n$. If $\mathcal{X}$ is equipped with the standard Euclidean norm, observe that J is the identity map. More generally, if $\mathcal{X}$ is equipped with an L^p-norm with $p > 1$, then J is single-valued and given for all $x \neq 0$ by

$$J(x) = \frac{\|x\|_p}{\|x\|_q} x \ ,$$

where q is the positive real number such that $\frac{1}{p} + \frac{1}{q} = 1$.

Example 4.10. Assume that the norm $\|\cdot\|$ on $\mathcal{X}$ is *polyhedral*, meaning that there is a finite symmetric family $\mathcal{W} \subset \mathcal{X}^*$ of linear forms on $\mathcal{X}$ such that

$$\|x\| = \max_{x^* \in \mathcal{W}} \langle x, x^* \rangle \ , \quad \forall x \in \mathcal{X} \ .$$

We may assume that $\mathcal{W}$ is minimal, in the sense that no element in $\mathcal{W}$ may be written as a convex combination of other elements in $\mathcal{W}$. This case includes the standard L^1- and L^∞-norms. Then, one may check that we have, for every $x \in \mathcal{X}$,

$$J(x) = \|x\| \ \operatorname{co}\{x^* \in \mathcal{W} \mid \langle x, x^* \rangle = \|x\|\} \ ,$$

where $\operatorname{co}(\mathcal{V})$ stands for the convex hull of any set $\mathcal{V}$.

Accretive maps

Accretive operators initially appeared as differential operators, and more precisely as infinitesimal generators of one-parameter semigroup of nonexpansive self-maps on a Banach space. We refer the reader to the series of papers [Bro67a, Bro67b, Bro68] by Browder and to the references therein for background on the subject, as well as to the more detailed monograph [Bro76]. See also Brezis [Bré73], Crandall and Liggett [CL71], Crandall and Pazy [CP72], Kato [Kat67], Martin [Mar70], for standard results in the theory of accretive operators.

Definition 4.11. Let $A : \mathcal{X} \rightrightarrows \mathcal{X}$ be a set-valued map. Denote by Id the identity map on $\mathcal{X}$.

- A is *accretive* if, for every $x, y \in \mathcal{X}$, every $u \in A(x)$ and every $v \in A(y)$, there exists an element $x^* \in J(x - y)$ such that $\langle u - v, x^* \rangle \geqslant 0$.
- A is *m-accretive* if it is accretive and if $\operatorname{rg}(\mathrm{Id} + \lambda A) = \mathcal{X}$ for every $\lambda > 0$.
- A is *coaccretive* if, for every $x, y \in \mathcal{X}$, every $u \in A(x)$ and every $v \in A(y)$, there exists an element $x^* \in J(u - v)$ such that $\langle x - y, x^* \rangle \geqslant 0$.

Let us mention that in a Hilbert space, the normalized duality mapping J is the identity map, and accretivity recovers the notion of *monotony*, while m-accretive maps are *maximal monotone* operators. In particular, note that the subdifferential of any lower semicontinuous proper convex function is m-accretive (see [RW98]).

Remark 4.12. An equivalent definition of accretivity is the following: a set-valued map $A : \mathcal{X} \rightrightarrows \mathcal{X}$ is accretive if, and only if, for every $x, y \in \mathcal{X}$, every $u \in A(x)$ and $v \in A(y)$, and every $\lambda > 0$, we have $\|x - y\| \leqslant \|x - y + \lambda(u - v)\|$. Furthermore, it is known that if $A : \mathcal{X} \rightrightarrows \mathcal{X}$ is accretive, then $\operatorname{rg}(\operatorname{Id} + \lambda A) = \mathcal{X}$ for *every* $\lambda > 0$ if, and only if, $\operatorname{rg}(\operatorname{Id} + \lambda A) = \mathcal{X}$ for *some* $\lambda > 0$. Finally, a set-valued map $A : \mathcal{X} \rightrightarrows \mathcal{X}$ is coaccretive if, and only if, A^{-1} is accretive.

4.3.2 Local boundedness of coaccretive maps

It is known that in a reflexive Banach space, any accretive map A is *locally bounded* at any point x in the interior of its domain, meaning that there exists a neighborhood $\mathcal{V}$ of x such that $A(\mathcal{V})$ is bounded, see [FHK72]. The following result shows that the same holds true for coaccretive maps, at least in finite dimension.

Proposition 4.13. *Let $(\mathcal{X}, \|\cdot\|)$ be a finite-dimensional real vector space, and let $A : \mathcal{X} \rightrightarrows \mathcal{X}$ be a coaccretive map. Then, for any vector x in the interior of* $\operatorname{dom}(A)$*, A is locally bounded at x.*

Proof. Assume by contradiction that A is not locally bounded at x, that is, for any neighborhood $\mathcal{V}$ of x, $A(\mathcal{V})$ is not bounded. Then, there exists a sequence $(x_k)_{k \in \mathbb{N}}$ in $\mathcal{X}$ converging to x such that, for all integers k, there is a vector $y_k \in A(x_k)$, and such that the sequence $(\|y_k\|)$ tends to infinity. Since the dimension is finite, the bounded sequence $(y_k / \|y_k\|)$ is converging, up to an extraction, toward some point $z \in \mathcal{X}$.

The point x being in the interior of $\operatorname{dom}(A)$, we may choose a scalar $\alpha > 0$ such that $x + \alpha z \in \operatorname{dom}(A)$. Let $y' \in A(x + \alpha z)$. Since A is coaccretive, for all integers k there is a linear form $y_k^* \in J(y_k - y')$ such that

$$\langle x_k - x - \alpha z, y_k^* \rangle \geqslant 0 \enspace .$$

By homogeneity of the duality mapping, we know that $z_k^* := \frac{y_k^*}{\|y_k\|} \in J\left(\frac{y_k - y'}{\|y_k\|}\right)$ for every integer k, and that

$$\langle x_k - x - \alpha z, z_k^* \rangle \geqslant 0 \enspace . \tag{4.2}$$

Since J is upper semicontinuous and with compact values (see Subsection 4.3.1), and since the sequence $((y_k - y')/\|y_k\|)$ converges, up to an extraction, toward z, we deduce that the sequence (z_k^*) has an accumulation point $z^* \in J(z)$. Then, Inequality (4.2) implies that $\langle z, z^* \rangle \leqslant 0$, a contradiction since $\langle z, z^* \rangle = \|z\|^2 = 1$ by definition of the duality mapping. ☐

We deduce an immediate corollary about the image of compact sets under coaccretive maps.

Corollary 4.14. *Let $(\mathcal{X}, \|\cdot\|)$ be a finite-dimensional real vector space, and let $A : \mathcal{X} \rightrightarrows \mathcal{X}$ be a coaccretive map. Then, for any compact set $\mathcal{K} \subset \mathcal{X}$ included in the interior of* $\operatorname{dom}(A)$*, $A(\mathcal{K})$ is bounded.*

4.3.3 Characterization of surjectivity of accretive maps

Let us denote by $\overline{B}(x, \rho)$ the closed ball centered at point $x \in \mathcal{X}$ with radius $\rho \geqslant 0$, and by $\operatorname{dist}_0(\mathcal{V})$ the distance of the origin to a subset $\mathcal{V} \subset \mathcal{X}$, i.e., $\operatorname{dist}_0(\mathcal{V}) := \inf_{x \in \mathcal{V}} \|x\|$.

To show the main theorem of this section, we shall need the following result. It is a special case of a result established by Kirk and Schöneberg [KS80] (see the last corollary of Th. 3) which is itself a generalization of theorems obtained by Lange [Lan71] and Kartsatos [Kar78, Kar81].

Theorem 4.15 ([KS80])**.** *Let $(\mathcal{X}, \|\cdot\|)$ be a finite-dimensional real vector space, and let $A : \mathcal{X} \rightrightarrows \mathcal{X}$ be an m-accretive set-valued map. Assume that*

$$\lim_{\|x\| \to \infty} \operatorname{dist}_0(A(x)) = +\infty \enspace .$$

Then, $\mathrm{rg}(A) = \mathcal{X}$.

We now state the main result of this section.

Theorem 4.16. *Let* $(\mathcal{X}, \|\cdot\|)$ *be a finite-dimensional real vector space, and let* $A : \mathcal{X} \rightrightarrows \mathcal{X}$ *be an accretive map. If* $\mathrm{rg}(A) = \mathcal{X}$ *then, for all scalars* $\alpha \geqslant 0$*, the subset* $\mathcal{D}_\alpha := \{x \in \mathcal{X} \mid \mathrm{dist}_0(A(x)) \leqslant \alpha\}$ *is bounded. Moreover, if* A *is* m*-accretive, then the converse is also true.*

Proof. Assume that $\mathrm{rg}(A) = \mathcal{X}$. Equivalently, we have $\mathrm{dom}(A^{-1}) = \mathcal{X}$. Since A is accretive, its inverse A^{-1} is coaccretive (see Subsection 4.3.1). Hence, according to Corollary 4.14, the image of any compact set under A^{-1} is bounded. Now, observe that for every nonnegative real numbers α, α' such that $\alpha' > \alpha$, we have $\mathcal{D}_\alpha \subset A^{-1}\left(\overline{B}(0,\alpha')\right)$. Indeed, $A^{-1}\left(\overline{B}(0,\alpha')\right) = \{x \in \mathcal{X} \mid A(x) \cap \overline{B}(0,\alpha') \neq \emptyset\}$ by definition. Hence, $\mathcal{D}_\alpha$ is bounded for every $\alpha \geqslant 0$.

For the converse property, let us assume that A is m-accretive. It is readily seen that the coercivity condition in Theorem 4.15 is equivalent to the boundedness of all subsets $\mathcal{D}_\alpha$, hence the result. □

Remark 4.17. Here, since the dimension is finite, all the norms are equivalent. Hence, in the latter theorem, the choice of the norm only matters for the definition of accretivity, but any other norm may be considered in the definition of dist_0.

Example 4.18 (*discrete* p*-Laplacian*)**.** Let $\mathcal{G} = (V, E)$ be a connected (undirected) graph with a finite set of nodes an no multiple edges. For brevity, we write $i \leftrightarrow j$ if there is an edge between i and j. Let us equip every edge $\{i,j\} \in E$ with a weight $C_{ij} > 0$. For $p > 1$, we denote by $L_p : \mathbb{R}^V \to \mathbb{R}^V$ the *discrete* p*-Laplacian*, whose ith entry is defined by

$$[L_p(v)]_i := \sum_{\substack{j \in V \\ i \leftrightarrow j}} C_{ij}\,(v_i - v_j)\,|C_{ij}\,(v_i - v_j)|^{p-2} \quad , \quad v \in \mathbb{R}^V \ .$$

Let B be a proper subset of V, i.e., different from $\emptyset$ and V. Assume that for every vertex $i \in B$, some real number w_i is given. We are interested in the following boundary value problem: given $g \in \mathbb{R}^{V \setminus B}$, find $v \in \mathbb{R}^V$ such that

$$\begin{cases} [L_p(v)]_i = -g_i \ , & \forall i \in V \setminus B \ , \\ v_i = w_i \ , & \forall i \in B \ . \end{cases} \tag{P$_g$}$$

Note that the discrete p-Laplacian L_p appears as the gradient of the energy function J_p defined on $\mathbb{R}^V$ by

$$J_p(v) := \sum_{\{i,j\} \in E} \frac{1}{pC_{ij}}\,|C_{ij}(v_i - v_j)|^p \quad , \quad v \in \mathbb{R}^V \ .$$

Then, Problem (P$_g$) is equivalent to the following optimization problem:

$$\text{minimize} \ \ J_p(v) + \sum_{i \in V \setminus B} g_i\, v_i \qquad \text{s.t.} \ \ v_i = w_i \ , \quad \forall i \in B \ .$$

In particular, for $p = 2$ it recovers the classical problem of computing the electrical potential on the graph $\mathcal{G}$, with a prescribed potential w_i at each node $i \in B$ and a prescribed current g_i passing through each node $i \in V \setminus B$, C_{ij} being the conductance of the edge $\{i,j\}$.

Let $\mathcal{X} = \mathbb{R}^{V \setminus B}$. We reformulate Problem (P$_g$) by introducing the operator $A : \mathcal{X} \to \mathcal{X}$ whose ith coordinate map is defined, for all $x \in \mathcal{X}$, by

$$A_i(x) := \sum_{\substack{j \in V \setminus B \\ i \leftrightarrow j}} C_{ij}\,(x_i - x_j)\,|C_{ij}\,(x_i - x_j)|^{p-2} + \sum_{\substack{j \in B \\ i \leftrightarrow j}} C_{ij}\,(x_i - w_j)\,|C_{ij}\,(x_i - w_j)|^{p-2} \ .$$

Hence, given $g \in \mathbb{R}^{V \setminus B}$, Problem ($\mathrm{P}_g$) is equivalent to finding a solution $x \in \mathcal{X}$ to the equation $A(x) = -g$. Since A is the gradient of a continuous convex function, then it is m-accretive, $\mathcal{X}$ being endowed with the standard Euclidean norm. Thus, according to Theorem 4.16, the latter equation has a solution for all $g \in \mathcal{X}$ if, and only if, the subsets $\mathcal{D}_\alpha = \{x \in \mathcal{X} \mid \|A(x)\|_\infty \leqslant \alpha\}$ are bounded for all $\alpha \geqslant 0$. We next show that this is the case when the graph $\mathcal{G}$ is connected.

Let $\alpha \geqslant 0$ be such that $\mathcal{D}_\alpha$ is not empty. Let $x \in \mathcal{D}_\alpha$ and choose $i_0 \in V \setminus B$ such that $|x_{i_0}| = \|x\|_\infty$. We may assume, without loss of generality, that $x_{i_0} \geqslant 0$. If $x_{i_0} \leqslant \max_{j \in B} w_j$, then x is bounded by a constant which only depends on $w \in \mathbb{R}^B$. Assume now that $x_{i_0} > \max_{j \in B} w_j$. Since $\mathcal{G}$ is connected, there exists an elementary path (i.e., without loop), $i_0, i_1, \dots, i_m$, connecting i_0 to some node $i_m \in B$ and such that $i_\ell \in V \setminus B$ for all $\ell \in \{0, 1, \dots, m-1\}$. By induction, we easily show that $|x_{i_{\ell-1}} - x_{i_\ell}|$ is bounded for all $\ell \in [m]$ by a constant M that only depends on α, p and the parameters of $\mathcal{G}$. Thus, we obtain that $\|x\|_\infty = x_{i_0} \leqslant mM + w_{i_m} \leqslant |V|M + \|w\|_\infty$, where $|V|$ is the cardinality of V. This shows that $\mathcal{D}_\alpha$ is bounded.

Hence, Problem (P_g) has solution for all $g \in \mathbb{R}^{V \setminus B}$.

4.4 Accretive operator approach – application to nonexpansive maps

4.4.1 Existence stability of fixed points under additive perturbations

Let $(\mathcal{X}, \|\cdot\|)$ be a finite-dimensional real vector space. Recall that any operator $T : \mathcal{X} \to \mathcal{X}$ is nonexpansive (with respect to $\|\cdot\|$) if $\|T(x) - T(y)\| \leqslant \|x - y\|$ for all $x, y \in \mathcal{X}$. The following lemma relates nonexpansive operators to accretive maps.

Lemma 4.19. *Let $T : \mathcal{X} \to \mathcal{X}$ be a nonexpansive operator. Then, the map $A := \mathrm{Id} - T$ is m-accretive.*

Proof. We first show that A is accretive. Let $x, y \in \mathcal{X}$, and let $x^* \in J(x - y)$. We have

$$\langle T(x) - T(y), x^* \rangle \;\leqslant\; \|x^*\|_* \, \|T(x) - T(y)\| \;\leqslant\; \|x^*\|_* \, \|x - y\| \;=\; \langle x - y, x^* \rangle \enspace .$$

We deduce that $\langle A(x) - A(y), x^* \rangle = \langle x - y - (T(x) - T(y)), x^* \rangle \geqslant 0$, which proves that A is accretive.

Now, let $\lambda > 0$ and $z \in \mathcal{X}$. The vector z is in the range of $\mathrm{Id} + \lambda A$ if, and only if, the map $x \mapsto T_{\lambda,z}(x) := \frac{\lambda}{1+\lambda} T(x) + \frac{1}{1+\lambda} z$ has a fixed point. Since T is nonexpansive, $T_{\lambda,z}$ is a contraction. More precisely, for all $x, y \in \mathcal{X}$, we have

$$\|T_{\lambda,z}(x) - T_{\lambda,z}(y)\| \leqslant \frac{\lambda}{1+\lambda} \|x - y\| \enspace ,$$

with $\frac{\lambda}{1+\lambda} < 1$. Hence, by the Banach fixed-point theorem, $T_{\lambda,z}$ has a (unique) fixed point. This proves that $\mathrm{rg}(\mathrm{Id} + \lambda A) = \mathcal{X}$ for every $\lambda > 0$, and so A is m-accretive. □

Using the previous lemma, we now adapt Theorem 4.16 to the case of nonexpansive maps. Thus, we get a necessary and sufficient condition for the existence of a fixed point for any additive perturbation of any nonexpansive operator.

Corollary 4.20. *Let $(\mathcal{X}, \|\cdot\|)$ be a finite-dimensional real vector space, and let $T : \mathcal{X} \to \mathcal{X}$ be a nonexpansive operator. Then, the following are equivalent:*

(i) for every vector $g \in \mathcal{X}$, the operator $g + T$ has a fixed point;

(ii) every nonexpansive operator $G : \mathcal{X} \to \mathcal{X}$ such that $\sup_{x \in \mathcal{X}} \|T(x) - G(x)\| < \infty$ has a fixed point;

(iii) for every scalar $\alpha \geqslant 0$, the set $\mathcal{D}_\alpha(T) := \{x \in \mathcal{X} \mid \|x - T(x)\| \leqslant \alpha\}$ is bounded.

Proof. The equivalence between Point (i) and Point (iii) is a mere application of Theorem 4.16 to $\mathrm{Id} - T$, which is m-accretive according to Lemma 4.19. Moreover, it is straightforward to check that (ii) $\Rightarrow$ (i).

We now prove that (i) $\Rightarrow$ (ii). Assume that (i) holds and let $G : \mathcal{X} \to \mathcal{X}$ be a nonexpansive operator such that $\sup_{x \in \mathcal{X}} \|T(x) - G(x)\| < \infty$. Since (i) $\Leftrightarrow$ (iii), we know that all the subsets $\mathcal{D}_\alpha(T)$ are bounded. We readily deduce that all the subsets $\mathcal{D}_\alpha(G)$ are also bounded. This implies in particular that G has a fixed point. □

4.4.2 Characterization of ergodicity of stochastic games

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a Shapley operator. Recall that T can be quotiented into a self-map $[T]$ of the additive projective space $\mathbb{TP}^n$. This quotiented map is nonexpansive with respect to the norm q_{H}, the quotiented version of Hilbert's seminorm on $\mathbb{R}^n$. Moreover, the solvability of the ergodic equation is equivalent to the existence of a fixed point for $[T]$ (see Subsection 2.3.4). These considerations allow us to apply the result of the previous subsection to characterize the ergodicity of stochastic games.

Theorem 4.21. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a Shapley operator. The following are equivalent:*

(i) T is ergodic, i.e., for all vectors $g \in \mathbb{R}^n$, the ergodic equation (2.5) is solvable for $g + T$;

(ii) for all scalars $\alpha \geqslant 0$, the space $\mathcal{D}_\alpha^{\mathrm{H}}(T) := \{x \in \mathbb{R}^n \mid \|x - T(x)\|_{\mathrm{H}} \leqslant \alpha\}$ is bounded in Hilbert's seminorm;

(iii) for all scalars $\alpha, \beta \in \mathbb{R}$, the slice space $\mathcal{S}_\alpha^\beta(T) := \{x \in \mathbb{R}^n \mid \alpha e + x \leqslant T(x) \leqslant \beta e + x\}$ is bounded in Hilbert's seminorm;

(iv) the ergodic equation (2.5) is solvable for all Shapley operators G on $\mathbb{R}^n$ such that $\sup_{x \in \mathbb{R}^n} \|T(x) - G(x)\|_{\mathrm{H}} < \infty$.

Proof. From the definition of Hilbert's seminorm, we see that $\mathcal{S}_\alpha^\beta(T) \subset \mathcal{D}_\delta^{\mathrm{H}}(T)$ as soon as $\beta - \alpha \leqslant \delta$. It readily follows that (ii) $\Rightarrow$ (iii).

Let $g \in \mathbb{R}^n$, and choose $\alpha, \beta \in \mathbb{R}$ such that $\alpha e \leqslant g + T(0) \leqslant \beta e$. Then, $\mathcal{S}_\alpha^\beta(g + T)$ is nonempty, and we know that it is invariant under $g + T$. Moreover, we have $\mathcal{S}_\alpha^\beta(g + T) \subset \mathcal{S}_{\alpha'}^{\beta'}(T)$, with $\alpha' := \alpha - \|g\|_\infty$ and $\beta' := \beta + \|g\|_\infty$. Thus, we deduce from Theorem 2.6 that (iii) $\Rightarrow$ (i).

Assume now that Point (i) holds. Then, in the quotient space $(\mathbb{TP}^n, q_{\mathrm{H}})$, the nonexpansive map $[g] + [T] = [g + T]$ has a fixed point for all equivalence classes $[g] \in \mathbb{TP}^n$. Applying Corollary 4.20, we get that for all scalars $\alpha \geqslant 0$, the space

$$\mathcal{D}_\alpha([T]) = \{[x] \in \mathbb{TP}^n \mid q_{\mathrm{H}}([x] - [T(x)]) \leqslant \alpha\}$$

is bounded (in the norm q_{H}). Since $x \in \mathcal{D}_\alpha^{\mathrm{H}}(T)$ if, and only if, $[x] \in \mathcal{D}_\alpha([T])$, we deduce that (i) $\Rightarrow$ (ii). Likewise, we readily get from Corollary 4.20 that (iv) $\Leftrightarrow$ (i). □

Chapter 5

Fixed-point problems for payment-free Shapley operators

In the previous chapter, a characterization of ergodicity for stochastic games with bounded payments (Assumption 4.A) has been given in terms of fixed points of some payment-free Shapley operator (Theorem 4.6). In this chapter, we first give a combinatorial characterization of the latter condition. This characterization involves a pair of directed hypergraphs. Then, we address some algorithmic and complexity issues related to the problem of checking ergodicty of a stochastic game and to the structure of the fixed-point set of a payment-free Shapley operator.

The results of this chapter are based on the paper [AGH14a], published in *Discrete and Continuous Dynamical Systems, Series A*.

5.1 Galois connection associated with payment-free Shapley operators

5.1.1 Preliminaries on lattices and Galois connections

Let $(L, \prec)$ be a partially ordered set (poset, for short), and let X be a subset of L. An element $a \in L$ is a *lower bound* for X if $a \prec x$ for all $x \in X$. The subset X has a *greatest lower bound*, or *infimum*, if there exists a lower bound a' for X such that $a \prec a'$ for every other lower bound a. Dually, an element $b \in L$ is an *upper bound* for X if $x \prec b$ for all $x \in X$. The subset X has a *least upper bound*, or *supremum*, if there exists an upper bound b' for X such that $b' \prec b$ for every other upper bound b. Furthermore, if $b' \in X$, then the latter element is said to be the maximum of X, and we denote it by $\max X$.

A poset L is an *inf-semilattice* (resp., *sup-semilattice*) is every nonempty finite subset has a greatest lower bound (resp., least upper bound). It is a *lattice* if it is both an inf-semilattice and a sup-semilattice. In this work, we shall consider lattices the elements of which are subsets, equipped with the inclusion partial order, and particularly the powerset lattice $\mathscr{P}(X)$ of any set X.

We next recall the definition of a Galois connection. This notion was first introduced by Birkhoff for lattices of subsets, in the first edition of [Bir79], and then generalized by Ore [Ore44]. Let $(A, \prec_A)$ and $(B, \prec_B)$ be two posets, and let $\phi : A \to B$ and $\psi : B \to A$ be two maps. The map ϕ is *antitone* if $a \prec_A a'$ implies $\phi(a') \prec_B \phi(a)$ (and similarly for ψ). The pair (ϕ, ψ) is a *Galois connection* between A and B if it satisfies one of the equivalent properties listed in the following theorem, where id_A (resp., id_B) denotes the identity map on A (resp., on B).

Theorem 5.1. *Let $(A, \prec_A)$ and $(B, \prec_B)$ be two posets, and let $\phi : A \to B$ and $\psi : B \to A$. The following are equivalent:*

(i) $\mathrm{id}_A \prec_A \psi \circ \phi$, $\mathrm{id}_B \prec_B \phi \circ \psi$, *and* ϕ, ψ *are antitone;*

(ii) $b \prec_B \phi(a) \iff a \prec_A \psi(b)$, $\forall a \in A$, $\forall b \in B$;

(iii) ϕ *is antitone and* $\psi(b) = \max\{a \in A \mid b \prec_B \phi(a)\}$, $\forall b \in B$;

(iv) ψ *is antitone and* $\phi(a) = \max\{b \in B \mid a \prec_A \psi(b)\}$, $\forall a \in A$.

Point (i) is the definition introduced by Birkhoff. The equivalence with Point (ii) is straightforward, and the proof of the equivalences with Points (iii) and (iv) may be found in [BJ72, Ch. 2] or [GHK+80, Ch. 0, Sec. 3].

It follows from Point (iii) of Theorem 5.1 that for any antitone map $\phi : A \to B$, there is at most one map $\psi : B \to A$ such that (ϕ, ψ) is a Galois connection between A and B. Let us denote this map by $\phi^\star := \psi$. The inequalities in Point (i) yields

$$\phi \circ \phi^\star \circ \phi = \phi \quad \text{and} \quad \phi^\star \circ \phi \circ \phi^\star = \phi^\star \ . \tag{5.1}$$

This implies in particular that, for all $b \in B$,

$$(\exists a \in A, \quad b = \varphi(a)) \iff \phi \circ \phi^\star(b) = b \ .$$

Also, observe that the definition of a Galois connection is symmetric, in the sense that (ϕ, ψ) is a Galois connection between A and B if, and only if, (ψ, ϕ) is a Galois connection between B and A. Hence, for any statement involving the Galois connection (ϕ, ψ), a dual statement holds for the Galois connection (ψ, ϕ). In particular, we get that

$$(\phi^\star)^\star = \phi \ .$$

We say that an element $a \in A$ is *closed* (with respect to the Galois connection $(\phi, \phi^\star)$) if $a = \phi^\star \circ \phi(a)$. We can show that the set of closed elements in A with respect to $(\phi, \phi^\star)$ is $\bar{A} = \phi^\star(B)$. Dually, the set of closed element in B is $\bar{B} := \{b \in B \mid b = \phi^\star \circ \phi(b)\} = \phi(A)$. Then, ϕ is a bijection from $\bar{A}$ to $\bar{B}$, and its inverse is $\phi^\star$.

5.1.2 Galois connection between invariant faces of the hypercube

Before introducing the Galois connection that will be useful to describe the fixed points of a payment-free Shapley operator, let us make some observation. We shall us in the remainder the following notation: the complement of any subset $I \subset [n]$ is denoted either by I^c or $[n] \setminus I$, and we denote by e_I the vector in $\mathbb{R}^n$ with entries 1 on I and 0 on I^c:

$$[e_I]_j = 1, \ \forall j \in I \quad \text{and} \quad [e_I]_j = 0, \ \forall j \notin I \ .$$

Also, for $x \in \mathbb{R}^n$, we denote

$$\arg\min x := \{i \in [n] \mid x_i = \min_{\ell \in [n]} x_\ell\} \quad \text{and} \quad \arg\max x := \{i \in [n] \mid x_i = \max_{\ell \in [n]} x_\ell\} \ .$$

Lemma 5.2. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a payment-free Shapley operator. If $u \in \mathbb{R}^n$ is a fixed point of F then, denoting by $I := \arg\min u$ and $J := \arg\max u$, we have*

$$F(e_{I^c}) \leqslant e_{I^c} \ , \tag{5.2}$$

$$e_J \leqslant F(e_J) \ . \tag{5.3}$$

Proof. Let us assume that $\min_{\ell\in[n]} u_\ell = 0$. Since F is additively homogeneous, there is no loss of generality. We also may assume that $e_{I^c} \leqslant u$, up to the product by a positive constant. Hence, by monotonicity of F, we get that $F(e_{I^c}) \leqslant u$. In particular, we have $F_i(e_{I^c}) \leqslant 0$ for every index $i \in I$. Moreover, since $e_{I^c} \leqslant e$, and since the unit vector is a fixed point of F, we have $F(e_{I^c}) \leqslant e$. It follows that $F(e_{I^c}) \leqslant e_{I^c}$.

We show the second inequality using the same arguments, making the nonrestrictive assumption that $\max_{\ell\in[n]} u_\ell = 1$ and $u \leqslant e_J$. □

Remark 5.3. Conditions (5.2) and (5.3) are dual. Indeed, let $\widetilde{F}$ be the conjugate operator of F defined by $\widetilde{F}(x) := -F(-x)$. Then, it is readily checked that $\widetilde{F}$ is a payment-free Shapley operator. Moreover, since $\widetilde{F}(x) = e - F(e - x)$, then $\widetilde{F}(e_I) = e - F(e_{I^c})$ for any subset I of $[n]$. Hence, condition (5.2) holds for F and I if, and only if, condition (5.3) holds for $\widetilde{F}$ and I. Furthermore, if u is a (nontrivial) fixed point of F, then the vector $\tilde{u} := e - u$ is a (nontrivial) fixed point of $\widetilde{F}$, verifying $\arg\max \tilde{u} = \arg\min u$.

Remark 5.4. The inequalities in Lemma 5.2 may be stated in geometric terms. First note that the monotonicity of F and the fact that every trivial vector (i.e., proportional to the unit vector) is a fixed point yields that the hypercube $[0,1]^n$ is invariant by F. Thus, it is readily seen that

$$F(e_{I^c}) \leqslant e_{I^c} \iff F_i(e_{I^c}) = 0 \ , \quad \forall i \in I \ , \tag{5.4}$$
$$e_J \leqslant F(e_J) \iff F_j(e_J) = 1 \ , \quad \forall j \in J \ . \tag{5.5}$$

Now, let us define

$$\mathcal{K}_I^- := \{x \in [0,1]^n \mid x_i = 0, \ \forall i \in I\} \quad \text{and} \quad \mathcal{K}_J^+ := \{x \in [0,1]^n \mid x_j = 1, \ \forall j \in J\} \ ,$$

two faces of the hypercube. We shall call them lower and upper faces, respectively. Alternatively, we can define these faces by

$$\mathcal{K}_I^- = \{x \in [0,1]^n \mid x \leqslant e_{I^c}\} \quad \text{and} \quad \mathcal{K}_J^+ = \{x \in [0,1]^n \mid x \geqslant e_J\} \ .$$

From the monotonicity of F, we deduce that the inequalities in Lemma 5.2 are equivalent to the invariance of such faces of the hypercube, that is,

$$F(e_{I^c}) \leqslant e_{I^c} \iff F(\mathcal{K}_I^-) \subset \mathcal{K}_I^- \ ,$$
$$e_J \leqslant F(e_J) \iff F(\mathcal{K}_J^+) \subset \mathcal{K}_J^+ \ .$$

Given a payment-free Shapley operator F on $\mathbb{R}^n$, we now introduce the two families of subsets of $[n]$, denoted by $\mathcal{F}^+$ and $\mathcal{F}^-$, satisfying the inequalities (5.2) and (5.3), respectively:

$$\mathcal{F}^+ := \{I \subset [n] \mid F(e_{I^c}) \leqslant e_{I^c}\} \ ,$$
$$\mathcal{F}^- := \{J \subset [n] \mid e_J \leqslant F(e_J)\} \ .$$

According to the geometric interpretation (Remark 5.4), the two families $\mathcal{F}^+$ and $\mathcal{F}^-$ can be identified with the families of lower and upper invariant faces of the hypercube, respectively. However, the choice of the superscripts is motivated by the fact that $I \in \mathcal{F}^+$ (resp., $J \in \mathcal{F}^-$) if the vector e_{I^c} (resp., e_J) is a "superharmonic" (resp., "subharmonic") vector of F.

The latter families are lattices of subset with respect to the inclusion partial order. Indeed, since F is monotone, for all $I_1, I_2 \in \mathcal{F}^+$, we have

$$\begin{aligned} F(e_{(I_1\cup I_2)^c}) &= F\left(\inf\{e_{(I_1)^c}, e_{(I_2)^c}\}\right) \\ &\leqslant \inf\left\{F(e_{(I_1)^c}), F(e_{(I_2)^c})\right\} \\ &\leqslant \inf\{e_{(I_1)^c}, e_{(I_2)^c}) = e_{(I_1\cup I_2)^c} \ , \end{aligned}$$

so that $I_1 \cup I_2 \in \mathcal{F}^+$. This implies that the supremum of two sets in $\mathcal{F}^+$ coincides with their supremum in the powerset lattice $\mathscr{P}([n])$, namely the union $I_1 \cup I_2$. Hence, $\mathcal{F}^+$ is a sup-subsemilattice of $\mathscr{P}([n])$. Moreover, $\mathcal{F}^+$ has a bottom element (namely the empty set) and it is a finite poset. Then, it is automatically an inf-semilattice. Check that the infimum of two sets $I_1, I_2 \in \mathcal{F}^+$ is given by

$$\bigcup_{\substack{I_3 \in \mathcal{F}^+ \\ I_3 \subset I_1,\, I_3 \subset I_2}} I_3 \ .$$

Note that the latter infimum may differ from the infimum in $\mathscr{P}([n])$, namely the intersection $I_1 \cap I_2$. We prove that $\mathcal{F}^-$ is a lattice by dual arguments. Also, note that $\mathcal{F}^+$ and $\mathcal{F}^-$ both contain $\emptyset$ and $[n]$.

Following the conclusions of Lemma 5.2, given a subset $I \in \mathcal{F}^+$, we are interested in the subsets $J \in \mathcal{F}^-$ satisfying $I \cap J = \emptyset$. We shall consider in particular the greatest subset J with the latter property. Vice versa, given a subset $J \in \mathcal{F}^-$, we shall consider the greatest subset $I \in \mathcal{F}^+$ such that $I \cap J = \emptyset$. Put in geometric terms, to each lower invariant face $\mathcal{K}_I^-$ of the hypercube $[0,1]^n$, we consider the greatest upper invariant face $\mathcal{K}_J^+$ with nonempty intersection with $\mathcal{K}_I^-$. Let us show that this defines a Galois connection between the lattices $\mathcal{F}^+$ and $\mathcal{F}^-$.

Let $(\Phi, \Phi^\star)$ be the pair of functions from $\mathcal{F}^+$ to $\mathcal{F}^-$ and from $\mathcal{F}^-$ to $\mathcal{F}^+$, respectively, that have just been introduced. Formally, they are defined for every $I \in \mathcal{F}^+$ and $J \in \mathcal{F}^+$ by:

$$\Phi(I) := \bigcup_{\substack{J \in \mathcal{F}^- \\ I \cap J = \emptyset}} J \qquad \text{and} \qquad \Phi^\star(J) := \bigcup_{\substack{I \in \mathcal{F}^+ \\ I \cap J = \emptyset}} I \ . \tag{5.6}$$

It follows readily from their definition that Φ and $\Phi^\star$ are antitone, that $I \subset \Phi^\star \circ \Phi(I)$ and $J \subset \Phi \circ \Phi^\star(J)$. Hence condition (i) in Theorem 5.1 is satisfied for the pair $(\Phi, \Phi^\star)$ which proves that it is a Galois connection between the lattices of subsets $\mathcal{F}^+$ and $\mathcal{F}^-$.

5.1.3 Galois connection and nontrivial fixed points

We now explore some properties of the Galois connection introduced in the previous subsection. Let us first rephrase and complete Lemma 5.2, by a direct application of its definition.

Lemma 5.5. *Let F be a payment-free Shapley operator on $\mathbb{R}^n$. If u is a fixed point of F, then*

$$\arg\min u \in \mathcal{F}^+ \quad \textit{and} \quad \arg\max u \in \mathcal{F}^- \ .$$

Furthermore, if u is nontrivial, we have

$$\arg\max u \subset \Phi(\arg\min u) \quad \textit{and} \quad \arg\min u \subset \Phi^\star(\arg\max u) \ .$$

□

In the next lemma, we use the notation

$$F^\omega(x) := \lim_{k \to \infty} F^k(x) \ , \quad x \in \mathbb{R}^n \ ,$$

as soon as the limit exists. This is the case in particular when $F(x) \leqslant x$ or $x \leqslant F(x)$. Indeed, since F is monotone, the former (resp., the latter) inequality implies that the sequence $(F^k(x))_{k \in \mathbb{N}}$ is nonincreasing (resp., nondecreasing). Moreover, since F is nonexpansive in the sup-norm and has a fixed point (namely 0), the sequence $(F^k(x))_{k \in \mathbb{N}}$ is bounded, so that it converges to a finite point when k tends to infinity. Also, observe that $F^\omega(x)$, if it is well defined, is a fixed point of F.

Lemma 5.6. *Let F be a payment-free Shapley operator on $\mathbb{R}^n$. Let $I \in \mathcal{F}^+$ and $J \in \mathcal{F}^-$ be such that $\Phi(I) \neq \emptyset$ and $\Phi^\star(J) \neq \emptyset$. Then,*

$$\arg\max F^\omega(e_{I^c}) = \Phi(I) \quad \textit{and} \quad \arg\min F^\omega(e_J) = \Phi^\star(J) \ .$$

Furthermore, if I (resp., J) is closed with respect to the Galois connection $(\Phi, \Phi^\star)$ (resp., $(\Phi^\star, \Phi)$), then

$$\arg\min F^\omega(e_{I^c}) = I \quad (\textit{resp.,} \ \arg\max F^\omega(e_J) = J) \ .$$

Proof. First, note that since $F(e_{I^c}) \leqslant e_{I^c}$, the sequence $(F^k(e_{I^c}))_{k\in\mathbb{N}}$ is nonincreasing and so the limit $u := F^\omega(e_{I^c})$ does exist. Since F leaves $[0,1]^n$ invariant, we have $u \in [0,1]^n$.

Second, by definition of the Galois connection, we have $e_{\Phi(I)} \leqslant e_{I^c}$. Using the monotonicity of F and the characterization of $\mathcal{F}^+$ and $\mathcal{F}^-$, we get that

$$e_{\Phi(I)} \leqslant F(e_{\Phi(I)}) \leqslant F(e_{I^c}) \leqslant e_{I^c} \ ,$$

and again, by monotonicity of F, we obtain that, for all integers k,

$$e_{\Phi(I)} \leqslant F^k(e_{\Phi(I)}) \leqslant F^k(e_{I^c}) \leqslant e_{I^c} \ .$$

It follows that $e_{\Phi(I)} \leqslant u \leqslant e_{I^c}$ and we deduce that $I \subset \arg\min u$ and $\Phi(I) \subset \arg\max u$. Since Φ is antitone, we readily have $\Phi(\arg\min u) \subset \Phi(I) \subset \arg\max u$.

The vector u being a fixed point of F, we know from Lemma 5.5 that $\arg\min u \in \mathcal{F}^+$, $\arg\max u \in \mathcal{F}^-$, and $\arg\max u \subset \Phi(\arg\min u)$. Hence, from the previous inclusions we deduce that $\Phi(\arg\min u) = \Phi(I) = \arg\max u$.

Suppose now that I is closed with respect to the Galois connection. This means that $\Phi^\star \circ \Phi(I) = I$. Then, from the previous equalities, we get that $\Phi^\star \circ \Phi(\arg\min u) = I$. This implies that $\arg\min u \subset I$ and since we already know that $I \subset \arg\min u$, we conclude that $I = \arg\min u$.

The statements involving $J \in \mathcal{F}^-$ follow by duality. □

We next characterize the existence of a nontrivial fixed point for any payment-free Shapley operator in terms of the Galois connection $(\Phi, \Phi^\star)$. We shall say that $I, J \subset [n]$ are *conjugate subsets* with respect to the Galois connection $(\Phi, \Phi^\star)$ if

$$\begin{cases} I \in \mathcal{F}^+ \setminus \emptyset \ , \\ J \in \mathcal{F}^- \setminus \emptyset \ , \\ J = \Phi(I) \ , \\ I = \Phi^\star(J) \ . \end{cases}$$

Theorem 5.7. *A payment-free Shapley operator $F : \mathbb{R}^n \to \mathbb{R}^n$ has a nontrivial fixed point if, and only if, there exists a pair of conjugate subsets with respect to the Galois connection $(\Phi, \Phi^\star)$.*

Proof. Assume first that F has a nontrivial fixed point u. Let $I := \arg\min u$ and $J := \arg\max u$. We know from Lemma 5.5 that $I \in \mathcal{F}^+$ and $J \in \mathcal{F}^-$. Since $\Phi \circ \Phi^\star \circ \Phi = \Phi$, the subsets $\Phi^\star \circ \Phi(I)$ and $\Phi(I)$ are conjugate if, and only if, they are nonempty. This is the case since $I \subset \Phi^\star \circ \Phi(I)$ (by definition of Galois connections), and $J \subset \Phi(I)$ (by Lemma 5.5).

Assume now that (I, J) is a pair of conjugate subsets. Then, $\Phi(I) = J$ is nonempty and I is closed with respect to the Galois connection $(\Phi, \Phi^\star)$, since $\Phi^\star \circ \Phi(I) = \Phi^\star(J) = I$. Thus, it readily follows from Lemma 5.6 that $u := F^\omega(e_{I^c})$ is fixed point of F verifying $\arg\min u = I$ and $\arg\max u = J$. □

We say that a subset of any set X is *proper* if it differs from the empty set and from X. Using the properties of the Galois connection, we get alternative (milder) conditions characterizing the existence of a nontrivial fixed point.

Corollary 5.8. *A payment-free Shapley operator $F : \mathbb{R}^n \to \mathbb{R}^n$ has a nontrivial fixed point if, and only if, one of the following equivalent assertions holds:*

(i) there exist nonempty disjoint subsets $I, J \subset [n]$ such that $I \in \mathcal{F}^+$ and $J \in \mathcal{F}^-$;

(ii) there exists a nonempty subset $I \in \mathcal{F}^+$ such that $\Phi(I) \neq \emptyset$;

(iii) there exists a nonempty subset $J \in \mathcal{F}^-$ such that $\Phi^\star(J) \neq \emptyset$;

(iv) there is a proper subset that is closed with respect to the Galois connection $(\Phi, \Phi^\star)$ or $(\Phi^\star, \Phi)$;

(v) there exists a pair of conjugate subsets with respect to the Galois connection $(\Phi, \Phi^\star)$.

Proof. Since the equivalence between the existence of a nontrivial fixed point and condition (v) has been stated in Theorem 5.7, we only need to show the equivalence of Points (i) to (v).

(i) $\Rightarrow$ (ii), (iii): If (i) holds, then by definition of the Galois connection, we have $J \subset \Phi(I)$ and $I \subset \Phi^\star(J)$. Thus $\Phi(I) \neq \emptyset$ and $\Phi^\star(J) \neq \emptyset$, which shows both (ii) and (iii).

(ii) $\Rightarrow$ (iv): Let I be a subset as in (ii), that is, $I \in \mathcal{F}^+$ is nonempty and $\Phi(I) \neq \emptyset$. We cannot have $\Phi(I) = [n]$, otherwise, this would implies that $I \subset \Phi^\star \circ \Phi(I) = \emptyset$. Hence, $\Phi(I)$ is proper. Moreover, we know that $\Phi(I)$ is closed with respect to the Galois connection $(\Phi^\star, \Phi)$, which shows (one case of) (iv).

(iii) $\Rightarrow$ (iv): Similarly if $J \in \mathcal{F}^-$ is nonempty and if $\Phi^\star(J) \neq \emptyset$, then $\Phi^\star(J)$ is proper and closed with respect to the Galois connection $(\Phi, \Phi^\star)$, which shows (iv).

(iv) $\Leftrightarrow$ (v): This follows readily from the definition of conjugate subsets.

(v) $\Rightarrow$ (i): This also readily follows from the definition of conjugate subsets and of the Galois connection $(\Phi, \Phi^\star)$. □

5.2 Hypergraph characterization of the Galois connection

In this section, we introduce directed hypergraphs, which shall be useful to characterize the Galois connection $(\Phi, \Phi^\star)$. In particular, we shall see that finding $\Phi(I)$ (resp., $\Phi^\star(J)$) for any given $I \in \mathcal{F}^+$ (resp., $J \in \mathcal{F}^-$), is equivalent to solving a reachability problem in a directed hypergraph. We refer the reader to [GLNP93, All14] for background on reachability problems in hypergraphs.

5.2.1 Preliminaries on hypergraphs

A *directed hypergraph* is a pair (N, E) where N is a set of *nodes* and E is a set of *(directed) hyperarcs*. A hyperarc e is an ordered pair $(\mathbf{t}(e), \mathbf{h}(e))$ of disjoint nonempty subsets of nodes; $\mathbf{t}(e)$ is the *tail* of e and $\mathbf{h}(e)$ is its *head*. For brevity, we shall write $\mathbf{t}$ and $\mathbf{h}$ instead of $\mathbf{t}(e)$ and $\mathbf{h}(e)$, respectively, if no confusion can be made about the edge e. When $\mathbf{t}$ and $\mathbf{h}$ are both of cardinality one, the hyperarc is said to be an arc, and when every hyperarc is an arc, the directed hypergraph becomes a directed graph.

In the following, the term hypergraph will always refer to a directed hypergraph. The *size* of a hypergraph $\mathcal{G} = (N, E)$ is defined as $\text{size}(\mathcal{G}) := |N| + \sum_{e \in E} |\mathbf{t}(e)| + |\mathbf{h}(e)|$, where $|X|$ denotes the cardinality of any set X.

Let $\mathcal{G} = (N, E)$ be a hypergraph. A *hyperpath* of length m from a set of nodes $I \subset N$ to a node $j \in N$ is a sequence of m hyperarcs $(\mathbf{t}_1, \mathbf{h}_1), \dots, (\mathbf{t}_m, \mathbf{h}_m)$, such that $\mathbf{t}_i \subset \cup_{k=0}^{i-1} \mathbf{h}_k$ for all $i = 1, \dots, m+1$, with the convention $\mathbf{h}_0 = I$ and $\mathbf{t}_{m+1} = \{j\}$. Then, we say that a node $j \in N$ is *reachable* from a set $I \subset N$ if there exists a hyperpath from I to j. Alternatively, the relation of reachability can be defined in a recursive way: a node j is reachable from the set I if either

$j \in I$ or there exists a hyperarc $(\mathbf{t}, \mathbf{h})$ such that $j \in \mathbf{h}$ and every node of $\mathbf{t}$ is reachable from the set I. A set J is said to be *reachable* from a set I if every node of J is reachable from I. We denote by $\mathrm{reach}(I, \mathcal{G})$ the set of reachable nodes from I in $\mathcal{G}$. Note that, by definition, we always have $I \subset \mathrm{reach}(I, \mathcal{G})$.

A subset $I \subset N$ is *invariant* in the hypergraph $\mathcal{G}$ if it contains every node that is reachable from itself, that is $\mathrm{reach}(I, \mathcal{G}) = I$. One readily checks that the set of nodes in N that are reachable from I is the smallest invariant set in the hypergraph $\mathcal{G}$ containing I. The following example illustrate the notion of reachability.

Example 5.9. Figure 5.1 shows a directed hypergraph $\mathcal{G}$ with 9 nodes, 4 hyperarcs and 2 arcs. Check that node 8 is reachable from subset $\{1, 2\}$ through the hyperpath

$$(\{1,2\},\{3\}),\ (\{2\},\{4\}),\ (\{3,4\},\{6\}),\ (\{1\},\{5\}),\ (\{5,6\},\{8\})\ .$$

But node 9 is not reachable from $\{1, 2\}$ since the latter subset does not have access to node 7. Also, we have here

$$\mathrm{reach}(\{1,2\},\mathcal{G}) = \{1,2,3,4,5,6,8\}\ .$$

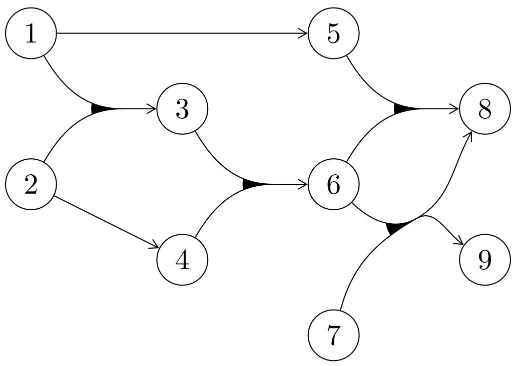

Figure 5.1: A directed hypergraph

5.2.2 Hypergraphs associated with payment-free Shapley operators

In order to characterize the Galois connection $(\Phi, \Phi^\star)$ associated with a payment-free Shapley operator F on $\mathbb{R}^n$, we introduce a pair of hypergraphs $(\mathcal{G}^+, \mathcal{G}^-)$ constructed as follows:

– the set of nodes for both $\mathcal{G}^+$ and $\mathcal{G}^-$ is $[n]$;
– the hyperarcs of $\mathcal{G}^+$ are the pairs $(J, \{i\})$ such that $F_i(e_J) > 0$ and $i \notin J$;
– the hyperarcs of $\mathcal{G}^-$ are the pairs $(J, \{i\})$ such that $F_i(e_{J^c}) < 1$ and $i \notin J$.

When F arises from a stochastic game, with a specific minimax representation, it is possible to give an interpretation of these hypergraphs in game-theoretic terms, providing a more intuitive understanding of the latter. These aspects shall be treated in Section 5.3. In the present section, we only exploit the properties (M), (AH) and (PH), characterizing payment-free Shapley operators.

The next proposition sheds light on the link between the hypergraphs $\mathcal{G}^\pm$ and the families of subsets $\mathcal{F}^\pm$. It is a direct consequence of the alternative characterizations (5.4) and (5.5) of the families $\mathcal{F}^\pm$, as well as the following observation. By monotonicity of F, if $(J, \{i\})$ is a hyperarc of $\mathcal{G}^\pm$, then $(J', \{i\})$ is also a hyperarc of $\mathcal{G}^\pm$ for any subset $J' \supset J$ which does not contain i. Hence, a subset $J \subset [n]$ is invariant in $\mathcal{G}^\pm$ if, and only if, there is no hyperarc in $\mathcal{G}^\pm$ from J to any node $i \notin J$.

Proposition 5.10. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a payment-free Shapley operator, to which is associated the pair of hypergraphs $(\mathcal{G}^+, \mathcal{G}^-)$. Then, a subset $I \subset [n]$ belongs to $\mathcal{F}^+$ (resp., $\mathcal{F}^-$) if, and only, if its complement in $[n]$ is invariant in $\mathcal{G}^+$ (resp., $\mathcal{G}^-$):*

$$I \in \mathcal{F}^+ \iff \operatorname{reach}(I^c, \mathcal{G}^+) = I^c \ ,$$
$$I \in \mathcal{F}^- \iff \operatorname{reach}(I^c, \mathcal{G}^-) = I^c \ .$$

□

We now make the link with the Galois connection $(\Phi, \Phi^\star)$.

Corollary 5.11. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a payment-free Shapley operator. Let $(\Phi, \Phi^\star)$ and $(\mathcal{G}^+, \mathcal{G}^-)$ be the Galois connection and the pair of hypergraphs, respectively, defined as above. Then, for all $I \in \mathcal{F}^+$ and $J \in \mathcal{F}^-$, we have*

$$\Phi(I) = [n] \setminus \operatorname{reach}(I, \mathcal{G}^-) \ ,$$
$$\Phi^\star(J) = [n] \setminus \operatorname{reach}(J, \mathcal{G}^+) \ .$$

Proof. It follows readily from the definition of Φ that $[n] \setminus \Phi(I)$ is the smallest subset I' containing I such that I'^c is in $\mathcal{F}^-$. By Proposition 5.10, the latter condition holds if, and only if, I' is invariant in $\mathcal{G}^-$. Hence, by definition of reachability, $\Phi(I)$ is the complement in $[n]$ of the set of nodes that are reachable from I in $\mathcal{G}^-$. For $\Phi^\star$ the argument is dual. □

We shall say that $I, J \subset [n]$ are *conjugate subsets* with respect to the hypergraphs $\mathcal{G}^\pm$, if

$$\begin{cases} I, J \neq \emptyset \ , \\ I = [n] \setminus \operatorname{reach}(J, \mathcal{G}^+) \ , \\ J = [n] \setminus \operatorname{reach}(I, \mathcal{G}^-) \ . \end{cases}$$

It readily follows from Corollary 5.11 that any sets $I, J \subset [n]$ are conjugate with respect to the Galois connection $(\Phi, \Phi^\star)$ if, and only if, they are conjugate with respect to the hypergraphs $(\mathcal{G}^+, \mathcal{G}^-)$. Hence, a reformulation of Corollary 5.8 in terms of the hypergraphs $(\mathcal{G}^+, \mathcal{G}^-)$ leads to the following.

Theorem 5.12. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a payment-free Shapley operator, to which is associated the pair of hypergraphs $(\mathcal{G}^+, \mathcal{G}^-)$. The following assertions are equivalent:*

(i) F has a nontrivial fixed point;
(ii) there exist nonempty disjoint subsets $I, J \subset [n]$ such that I^c and J^c are invariant in $\mathcal{G}^+$ and $\mathcal{G}^-$, respectively.
(iii) there exists a pair of conjugate subsets $I, J \subset [n]$ with respect to the hypergraphs $(\mathcal{G}^+, \mathcal{G}^-)$.

□

5.2.3 Ergodicity of stochastic games with bounded payments

As an immediate corollary of Theorem 5.12, we get a combinatorial characterization of the ergodicity of stochastic games with bounded payment function, relying on the analytical characterization of Theorem 4.6.

Corollary 5.13. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a Shapley operator satisfying Assumption 4.A, and let $(\mathcal{G}^+, \mathcal{G}^-)$ be the pair of hypergraphs associated with the recession operator of T. The following assertions are equivalent:*

(i) T is ergodic, i.e., for all vectors $g \in \mathbb{R}^n$, the ergodic equation (2.5) is solvable for $g + T$;

(ii) there do not exist nonempty disjoint subsets $I, J \subset [n]$ such that I^c and J^c are invariant in $\mathcal{G}^+$ and $\mathcal{G}^-$, respectively;

(iii) there does not exist conjugate subsets $I, J \subset [n]$ with respect to the hypergraphs $(\mathcal{G}^+, \mathcal{G}^-)$.

Remark 5.14. It is instructive to specialize the latter result to the case in which T arises as the dynamic programming operator of a finite Markov chain with rewards. Then, we can write $T(x) = g + Px$, where $g \in \mathbb{R}^n$ is the payment vector and P is a $n \times n$ stochastic matrix. We also have $\widehat{T}(x) = Px$.

In that case, the two hypergraphs $\mathcal{G}^+$ and $\mathcal{G}^-$ are identical. More precisely, there exists a hyperarc $(J, \{i\})$ in $\mathcal{G}^\pm$ if, and only if, there is some state $j \in J$ such that $P_{ij} > 0$. Consequently, the reachability relation in $\mathcal{G}^\pm$ is the same as in the directed graph $\mathcal{G}'$ with same set of nodes $[n]$ and an arc from j to i if $P_{ij} > 0$. Note that here, on contrary to hypergraphs, loops are possible. The latter directed graph $\mathcal{G}'$ is the transpose of the directed graph $\mathcal{G}$ associated with the transition matrix P, i.e., the arcs of $\mathcal{G}'$ are in the opposite direction. In particular, I^c is invariant in $\mathcal{G}'$ (hence in $\mathcal{G}^\pm$) if, and only if, there are no arcs nor paths from I to I^c in $\mathcal{G}$. Similarly, (I, J) is a pair of conjugate subsets with respect to $(\mathcal{G}^+, \mathcal{G}^-)$ if, and only if, I is the greatest set of nodes with no paths in $\mathcal{G}$ to a node of J, and vice versa.

So, Assertion (ii) of Corollary 5.13 implies the existence of a unique final class in $\mathcal{G}$. Indeed, if this is not the case, that is, if there exist two distinct final classes I, J in $\mathcal{G}$, then there are no arcs nor paths from I to I^c in $\mathcal{G}$, and the same is true for J, a contradiction. Conversely, if $\mathcal{G}$ has a unique final class, then Point (ii) of Corollary 5.13 is necessarily true. Also by contradiction, if there exist two nonempty disjoint subsets I, J such that there are no arcs from I to I^c, nor from J to J^c in $\mathcal{G}$, then there exists a final class of $\mathcal{G}$ included in I and also a final class of $\mathcal{G}$ included in J, a contradiction since I and J are disjoint. Hence in the present case, Assertion (ii) of Corollary 5.13 is equivalent to the condition that the directed graph associated with P has a unique final class, that is, Assertion (iv) in Theorem 4.1.

5.3 Game-theoretic interpretation

In this section, we consider a stochastic game $\Gamma = ([n], A, B, K_A, K_B, r, p)$, which satisfies the following standing assumption.

Assumption 5.B.
(i) The action spaces A and B are compact sets endowed with their respective Borel σ-algebra, and for all states $i \in [n]$, the sets of admissible actions A_i and B_i are nonempty closed sets.
(ii) The payment function r and the transition function p are continuous.

Note that with the latter assumption, the value in finite horizon and the discounted value exist. In particular, the Shapley operator, $T : \mathbb{R}^n \to \mathbb{R}^n$, of Γ is well defined and its ith component is given by

$$T_i(x) = \min_{\mu \in \Delta(A_i)} \max_{\nu \in \Delta(B_i)} \Big(r(i, \mu, \nu) + \sum_{j \in [n]} x_j \, p(j \mid i, \mu, \nu) \Big)$$
$$= \max_{\nu \in \Delta(B_i)} \min_{\mu \in \Delta(A_i)} \Big(r(i, \mu, \nu) + \sum_{j \in [n]} x_j \, p(j \mid i, \mu, \nu) \Big) \ , \quad x \in \mathbb{R}^n \ .$$

Let us make some comments. First, observe that inf and sup operators in (2.1) have been replaced by min and max operators, respectively. This is justified by the fact that the set of probability measures on A_i and B_i, respectively, are compact with respect to the weak-* topology, and the multilinear extensions $r(i, \cdot, \cdot)$ and $p(j \mid i, \cdot, \cdot)$ are continuous with respect to this topology. Another implication of Assumption 5.B is that the payment function is bounded. As a consequence, the recession operator of T exists and its ith component is given by

$$\widehat{T}_i(x) = \min_{\mu \in \Delta(A_i)} \max_{\nu \in \Delta(B_i)} \sum_{j \in [n]} x_j \, p(j \mid i, \mu, \nu) \ , \quad x \in \mathbb{R}^n \ ,$$

where the minimum and the maximum commutes. Moreover, Assumption 4.A is satisfied. Hence, according to Theorem 4.6, the ergodicity of Γ can be characterized by the absence of nontrivial fixed points for the payment-free Shapley operator $\widehat{T}$.

5.3.1 Dominions and ergodicity

We shall say that the families $\mathcal{F}^+$ and $\mathcal{F}^-$ of subsets of $[n]$ and the hypergraphs $\mathcal{G}^+$ and $\mathcal{G}^-$ are associated with the game Γ when the latter objects are constructed from the recession operator $\widehat{T}$. We next give a game-theoretic interpretation of the results obtained in the previous sections.

We shall call *dominion* of player MIN (resp., player MAX) in the game Γ any nonempty subset of states $I \subset [n]$ subject to the control of MIN (resp., MAX), in the sense that he has a pure stationary strategy such that a sequence of states starting in any state $i \in I$ stays in I almost surely, whatever the strategy of the other player is.

Proposition 5.15. *Let Γ be a stochastic game with finite state space satisfying Assumption 5.B, and let $(\mathcal{F}^+, \mathcal{F}^-)$ be the families of subsets of states associated with Γ. Then, $I \in \mathcal{F}^+ \setminus \emptyset$ if, and only if, I is a dominion of* MIN *in Γ. Likewise, $J \in \mathcal{F}^- \setminus \emptyset$ if, and only if, J is a dominion of* MAX *in Γ.*

Before giving the proof, let us recall the following definition: for any topological space X, endowed with its Borel σ-algebra, denoted by $\mathscr{B}$, the support of a measure μ on $(X, \mathscr{B})$ is the unique closed set, denoted by $\operatorname{supp} \mu$, satisfying:

(i) $\mu(X \setminus \operatorname{supp} \mu) = 0$;

(ii) If $U \subset X$ is open and $U \cap \operatorname{supp} \mu \neq \emptyset$, then $\mu(U \cap \operatorname{supp} \mu) > 0$.

When $X = [n]$, we shall identify a probability measure $\mu \in \Delta(X)$ with the stochastic vector $p \in \mathbb{R}^n$ whose ith entry is equal to $p_i = \mu(\{i\})$. The support of p is then given by $\operatorname{supp} p := \{i \in [n] \mid p_i > 0\}$.

Proof. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of Γ. Let $I \in \mathcal{F}^+ \setminus \emptyset$. Then, according to (5.4), we have $\widehat{T}_i(e_{I^c}) = 0$ for all $i \in I$. Let $\mu \in \Delta(A_i)$ be a mixed action of player MIN in state i attaining the minimum in the minimax formula of $\widehat{T}(e_{I^c})$. For all mixed actions $\nu \in \Delta(B_i)$ of player MAX, we have

$$\sum_{j \notin I} p(j \mid i, \mu, \nu) = 0 \tag{5.7}$$

By definition of the support of a measure, we necessarily have $\operatorname{supp} \mu \neq \emptyset$, otherwise $\mu(A_i) = 0$. So, we may choose an action $a_i \in A_i$ contained in $\operatorname{supp} \mu$. Then, it satisfies $p(I^c \mid i, a_i, b) = 0$ for all actions $b \in B_i$. Indeed, if this is not the case, then there exists some $b \in B_i$ such that $p(I^c \mid i, a_i, b)$ is bounded below by a positive constant. By continuity of $a \mapsto p(I^c \mid i, a, b)$, this implies that the function $a \mapsto p(I^c \mid i, a, b)$ is also bounded below by a positive constant on a neighborhood U of a_i, and since $\mu(U \cap \operatorname{supp} \mu) > 0$, this is a contradiction with (5.7).

Let σ be a pure stationary strategy of player MIN, identified with a map from $[n]$ to A, such that $\sigma(i) = a_i$ for all state $i \in I$. Suppose MIN chooses this strategy. At each step k of Γ, if the current state i_k is in I, and if player MAX chooses action b, then the probability that the state i_{k+1} at the following stage is in I^c is equal to $p(I^c \mid i_k, a_{i_k}, b) = 0$, whatever MAX chooses. Hence, if the initial state is in I, then for any strategy of MAX, the probability that the sequence of states $(i_k)_{k\geqslant 1}$ leaves I is 0. This shows the "only if" part of the first statement.

Conversely, suppose that there exists a pure stationary strategy σ of player MIN, mapping $i \in [n]$ to $a_i = \sigma(i) \in A_i$, such that for any initial state i_1 in I and any strategy of player MAX, the state of Γ stays in I almost surely. In particular, for any state $i \in I$ and any action $b \in B_i$, if the initial state is $i_1 = i$, if MIN chooses strategy σ and if MAX chooses action b at first stage, then the probability that i_2 is outside I is equal to 0. This probability coincides with $p(I^c \mid i, a_i, b)$. Hence, for all states $i \in [n]$, we have

$$\widehat{T}_i(e_{I^c}) \leqslant \max_{\nu \in \Delta(B_i)} \sum_{j \notin I} p(j \mid i, a_i, \nu) = 0 \enspace .$$

This shows that $\widehat{T}_i(e_{I^c}) = 0$ for all $i \in I$, that is $I \in \mathcal{F}^+$. The second assertion follows by duality. □

We deduce from the above proposition, combined with Corollary 5.8 and Theorem 4.6, a necessary and sufficient condition for ergodicity of stochastic games in terms of dominions

Corollary 5.16. *A stochastic game Γ with finite state space satisfying Assumption 5.B is ergodic if, and only if, there does not exist a pair of disjoint dominions of players* MIN *and* MAX*, respectively, in Γ.*

Remark 5.17. The hypergraphs $\mathcal{G}^+$ and $\mathcal{G}^-$ can also be interpreted in terms of strategies. Indeed, in $\mathcal{G}^-$, there is a hyperarc from a subset $J \in [n]$ to a state i if, and only if,

$$\min_{\mu \in \Delta(A_i)} \max_{\nu \in \Delta(B_i)} p(J^c \mid i, \mu, \nu) < 1 \enspace ,$$

which amounts to saying that in state i, player MIN has a mixed action μ_i such that, for any action b chosen by player MAX, the probability for the state to be in J at the next stage, i.e., $p(J \mid i, \mu_i, b)$, is bounded below by a positive constant. Note that, contrary to the family $\mathcal{F}^-$, we need to consider mixed actions of MIN (and not only pure actions).

The same interpretation holds for $\mathcal{G}^+$ and player MAX: there is a hyperarc from $J \subset [n]$ to $i \in [n]$ in $\mathcal{G}^+$ if, and only if, in state i, MAX has a mixed action such that, for any action chosen by MIN, the state at the next stage is in J with positive probability.

5.3.2 Ergodicity is a structural property

We call *upper* and *lower Boolean abstractions* of $\widehat{T}$, the operators defined on the set of Boolean vectors $\{0,1\}^n$, the ith component of which are given respectively by

$$\widehat{T}_i^+(x) := \min_{a \in A_i} \max_{b \in B_i} \max\{x_j \mid p(j \mid i, a, b) > 0\} \enspace ,$$

$$\widehat{T}_i^-(x) := \max_{b \in B_i} \min_{a \in A_i} \min\{x_j \mid p(j \mid i, a, b) > 0\} \enspace .$$

Let us make some observations. First, these Boolean operators can be extended to $\mathbb{R}^n$. Then, we have $\widehat{T}^- \leqslant \widehat{T} \leqslant \widehat{T}^+$. Second, the operators $\widehat{T}^+$ and $\widehat{T}^-$ are only determined by the supports of the transition probabilities, that is, by the set of all (i, a, b, j) such that $p(j \mid i, a, b) > 0$.

The latter Boolean operators are helpful to determine the set of dominions, as well as to construct the hypergraphs $\mathcal{G}^+$ and $\mathcal{G}^-$. These facts are a consequence of the following lemmas.

Lemma 5.18. *Let Γ be a stochastic game with state space $[n]$ satisfying Assumption 5.B, and with Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$. Then, a nonempty subset $I \subset [n]$ is a dominion of player* MIN *in Γ (i.e., $\widehat{T}(e_{I^c}) \leqslant e_{I^c}$) if, and only if, $\widehat{T}^+(e_{I^c}) \leqslant e_{I^c}$. Likewise, a nonempty subset $J \subset [n]$ is a dominion of player* MAX *in Γ (i.e., $e_J \leqslant \widehat{T}(e_J)$) if, and only if, $e_J \leqslant \widehat{T}^-(e_J)$.*

Proof. Recall that a nonempty subset I is a dominion of MIN if, and only if, $\widehat{T}(e_{I^c}) \leqslant e_{I^c}$ holds (Proposition 5.15). Furthermore, since $\widehat{T} \leqslant \widehat{T}^+$, we only need to prove the implication

$$\widehat{T}(e_{I^c}) \leqslant e_{I^c} \implies \widehat{T}^+(e_{I^c}) \leqslant e_{I^c} \ .$$

So, let us assume that $\widehat{T}(e_{I^c}) \leqslant e_{I^c}$, that is, I is a dominion of MIN. Then, for all $i \in I$, there exists an action $a_i \in A_i$ such that the probability $p(I^c \mid i, a_i, b)$ is equal to 0 for all $b \in B_i$. Let $x = e_{I^c}$. For all $i \in I$ and $b \in B_i$, we have $p(I^c \mid i, a_i, b) = \sum_j x_j \, p(j \mid i, a_i, b) = 0$. Since all the terms in the latter sum are nonnegative, we must have $x_j = 0$ as soon as $p(j \mid i, a_i, b) > 0$, which also writes $\max\{x_j \mid p(j \mid i, a_i, b) > 0\} = 0$. We deduce that $\widehat{T}_i^+(x) = 0$ for all $i \in I$, that is, $\widehat{T}^+(e_{I^c}) \leqslant e_{I^c}$. This shows the first equivalence, and the second follows by duality. □

Note that, according to the equivalences (5.4) and (5.5) and their Boolean counterparts, the equivalences in the previous lemma can be restated in the following way:

$$\widehat{T}_i(e_J) > 0 \iff \widehat{T}_i^+(e_J) = 1 \ ,$$
$$\widehat{T}_i(e_{J^c}) < 1 \iff \widehat{T}_i^-(e_{J^c}) = 0 \ ,$$

for all $i \in [n]$ and $J \subset [n]$. Hence, the following characterization.

Lemma 5.19. *Let Γ be a stochastic game with state space $[n]$ satisfying Assumption 5.B and with Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$. Let $\mathcal{G}^+$ and $\mathcal{G}^-$ be the hypergraphs associated with Γ. Then $(J, \{i\})$ is a hyperarc of $\mathcal{G}^+$ if, and only if, $\widehat{T}_i^+(e_J) = 1$ and $i \notin J$. Likewise, $(J, \{i\})$ is a hyperarc of $\mathcal{G}^-$ if, and only if, $\widehat{T}_i^-(e_{J^c}) = 0$ and $i \notin J$.* □

We deduce from the previous results that the set of dominions and the hypergraphs $\mathcal{G}^+$ and $\mathcal{G}^-$ are structural, in the sense that they only depend on the supports of the transition probabilities. Together with Corollary 5.16, we get in particular the following.

Corollary 5.20. *The ergodicity of a stochastic game Γ with finite state space satisfying Assumption 5.B only depends on the support of the transition probabilities.*

We conclude this section with the following theorem, providing a way to compute the images of subsets of states by the Galois connection $(\Phi, \Phi^\star)$.

Theorem 5.21. *Let Γ be a stochastic game with state space $[n]$ satisfying Assumption 5.B and with Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$. Let $\mathcal{F}^+$ and $\mathcal{F}^-$ be the families of subsets of states associated with Γ. Then, for $I \in \mathcal{F}^-$ and $J \in \mathcal{F}^+$ we have*

$$e_{\Phi(I)} = \big(\widehat{T}^-\big)^\omega(e_{I^c}) \ ,$$
$$e_{[n]\setminus\Phi^\star(J)} = \big(\widehat{T}^+\big)^\omega(e_J) \ .$$

Proof. We only show the first equality, the second follows by duality. Let $I \in \mathcal{F}^-$. We have $\widehat{T}^+(e_{I^c}) \leqslant e_{I^c}$, and using $\widehat{T}^- \leqslant \widehat{T}^+$, we obtain that $\widehat{T}^-(e_{I^c}) \leqslant e_{I^c}$. It follows that $(\widehat{T}^-)^\omega(e_{I^c})$ is well defined and satisfies $(\widehat{T}^-)^\omega(e_{I^c}) \leqslant e_{I^c}$. Since $\widehat{T}^-$ is a Boolean map, there exists some subset $L \subset [n]$ such that $e_L = (\widehat{T}^-)^\omega(e_{I^c})$. We need to show that $L = \Phi(I)$.

Since e_L is a fixed point of $\widehat{T}^-$, L belongs to $\mathcal{F}^+$. Furthermore, it satisfies $e_L \leqslant e_{I^c}$, since $I \cap L = \emptyset$. Then $L \subset \Phi(I)$.

Let $K \in \mathcal{F}^+$ be a subset such that $I \cap K = \emptyset$, that is, $e_K \leqslant e_{I^c}$. By induction, we get that $(\widehat{T}^-)^k(e_K) \leqslant (\widehat{T}^-)^k(e_{I^c})$ for every integer k. We also have $e_K \leqslant \widehat{T}^-(e_K)$, hence $(\widehat{T}^-)^\omega(e_K)$ exists and $(\widehat{T}^-)^\omega(e_K) \geqslant e_K$. This leads to $e_K \leqslant (\widehat{T}^-)^\omega(e_K) \leqslant (\widehat{T}^-)^\omega(e_{I^c}) = e_L$, which implies that $K \subset L$. This holds for all $K \in \mathcal{F}^+$ such that $I \cap K = \emptyset$, hence, by definition of Φ, we have $\Phi(I) \subset L$. □

5.4 Algorithmic and complexity issues

In this section, we consider a finite stochastic game $\Gamma = ([n], A, B, K_A, K_B, r, p)$ with perfect information. Assumption 5.B (hence, Assumption 4.A) is trivially satisfied, since the action spaces are finite. Moreover, the Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ of Γ is a polyhedral map, the ith component of which writes

$$T_i(x) = \min_{a \in A_i} \max_{b \in B_{i,a}} \Big(r(i,a,b) + \sum_{j \in [n]} x_j \, p(j \mid i,a,b) \Big) \enspace , \quad x \in \mathbb{R}^n \enspace .$$

As for its recession operator, which we denote here by F for simplicity, its ith component is given by

$$F_i(x) := \widehat{T}_i(x) = \min_{a \in A_i} \max_{b \in B_{i,a}} \sum_{j \in [n]} x_j \, p(j \mid i,a,b) \enspace , \quad x \in \mathbb{R}^n \enspace . \tag{5.8}$$

Note that all the results of Section 5.3 readily apply to the present setting, with minor adaptations coming from the fact that the min and max operators cannot be interchanged, see [AGH15a] where the perfect-information case has been specifically treated. In particular, the ith components of the Boolean abstractions of F are given by

$$F_i^+(x) = \widehat{T}_i^+(x) := \min_{a \in A_i} \max_{b \in B_{i,a}} \max\{x_j \mid p(j \mid i,a,b) > 0\} \enspace ,$$
$$F_i^-(x) = \widehat{T}_i^-(x) := \min_{a \in A_i} \max_{b \in B_{i,a}} \min\{x_j \mid p(j \mid i,a,b) > 0\} \enspace .$$

5.4.1 Syntactic hypergraphs

In this subsection, we introduce a pair of hypergraphs $(\mathcal{G}_s^+, \mathcal{G}_s^-)$ representing the Boolean operators F^+ and F^-. We shall see that the links between the Galois connection $(\Phi, \Phi^\star)$ and the hypergraphs $(\mathcal{G}^+, \mathcal{G}^-)$ established in Section 5.2 still hold with this new pair of hypergraphs. Hence, finding $\Phi(I)$ (resp., $\Phi^\star(J)$) for any subset $I \in \mathcal{F}^+$ (resp., $J \in \mathcal{F}^-$) is equivalent to solving a reachability problem in (a modified version of) $\mathcal{G}_s^-$ (resp., $\mathcal{G}_s^+$).

We first construct the hypergraph $\mathcal{G}_s^+$, a representation of which is shown in Figure 5.2. Let us introduce to that purpose a copy of $[n]$, which is a set disjoint from $[n]$ and given by a bijection c from $[n]$ to this copy. We also need to assume $\mathsf{c}([n])$ disjoint from the two constraints sets K_A and K_B.

The set of nodes of $\mathcal{G}_s^+$ is $[n] \cup K_A \cup \mathsf{c}([n])$. The hyperarcs of $\mathcal{G}_s^+$ are of the form:

– $(\{i\} \times A_i, \{\mathsf{c}(i)\})$ for each state $i \in [n]$;
– $(\{j\}, \{(i,a)\})$ for each states $i, j \in [n]$ and each action $a \in A_i$ such that there exists some action $b \in B_{i,a}$ with $p(j \mid i,a,b) > 0$.

As shown on Figure 5.2, this hypergraph is structured in two layers; the first layer consists of the arcs $(\{j\}, \{(i,a)\})$ whereas the second layer consists of the hyperarcs $(\{i\} \times A_i, \{\mathsf{c}(i)\})$. Furthermore, the size of $\mathcal{G}_s^+$ satisfies $\operatorname{size}(\mathcal{G}_s^+) = O(n^2|A|)$.

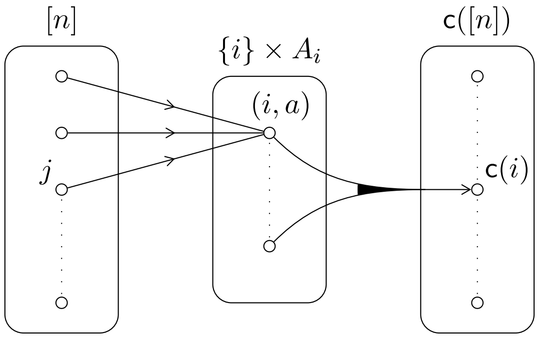

Figure 5.2: Hypergraph $\mathcal{G}_s^+$ associated with F^+

The hypergraph $\mathcal{G}_s^+$ encodes the Boolean operator F^+ in the following sense. Let $x \in \{0,1\}^n$ be a Boolean vector. Denote

$$x_{i,a} := \max_{b \in B_{i,a}} \max\{x_j \mid p(j \mid i, a, b) > 0\} \ ,$$

so that have

$$F_i^+(x) = \min_{a \in A_i} x_{i,a} \ .$$

If $x = e_J$ for any subset $J \subset [n]$, then $x_{i,a} = 1$ if, and only if, there exists some action $b \in B_{i,a}$ and some state $j \in J$ such that $p(j \mid i, a, b) > 0$. This is also equivalent to the node (i, a) being reachable from J in $\mathcal{G}_s^+$. Then, $F_i^+(x) = 1$ if, and only if, $x_{i,a} = 1$ for every action $a \in A_i$, which is equivalent to all the nodes in the tail of the hyperarc $(\{i\} \times A_i, \{\mathsf{c}(i)\})$ being reachable from J in $\mathcal{G}_s^+$. According to the recursive definition of reachability, this is equivalent to $\mathsf{c}(i)$ being reachable from J in $\mathcal{G}_s^+$. Hence, the following result.

Proposition 5.22. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be the payment-free Shapley operator associated with a perfect-information finite stochastic game. Then, the node $\mathsf{c}(i) \in \mathsf{c}([n])$ is reachable from the subset $J \subset [n]$ in $\mathcal{G}_s^+$ if, and only if, $F_i^+(e_J) = 1$.* □

We now construct the second hypergraph, a representation of which is shown in Figure 5.3. The node set of $\mathcal{G}_s^-$ is $[n] \cup K_B \cup \mathsf{c}([n])$, and its hyperarcs are:
- $(\{(i,a)\} \times B_i, \{\mathsf{c}(i)\})$ for each state $i \in [n]$ and action $a \in A_i$;
- $(\{j\}, \{(i,a,b)\})$ for each states $i, j \in [n]$ and each actions $a \in A_i$ and $b \in B_{i,a}$ such that $p(j \mid i, a, b) > 0$.

Again, the hypergraph $\mathcal{G}_s^-$ consists of two layers (see Figure 5.3). Furthermore, the size of $\mathcal{G}_s^-$ satisfies $\mathrm{size}(\mathcal{G}_s^-) = O(n^2|A||B|)$.

Like $\mathcal{G}_s^+$, the hypergraph $\mathcal{G}_s^-$ encodes the Boolean operator F^-, as shown by the following result.

Proposition 5.23. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be the payment-free Shapley operator associated with a perfect-information finite stochastic game. Then, the node $\mathsf{c}(i) \in \mathsf{c}([n])$ is reachable from the subset $J \subset [n]$ in $\mathcal{G}_s^-$ if, and only if, $F_i^-(e_{J^c}) = 0$.* □

Note that the absence of symmetry between $\mathcal{G}_s^+$ and $\mathcal{G}_s^-$ reflects the lack of symmetry between F^+ and F^-. Denote by $\overline{\mathcal{G}}_s^+$ and $\overline{\mathcal{G}}_s^-$ the hypergraphs obtained from $\mathcal{G}_s^+$ and $\mathcal{G}_s^-$, respectively, by identifying every node $i \in [n]$ with node $\mathsf{c}(i) \in \mathsf{c}([n])$. It follows readily from

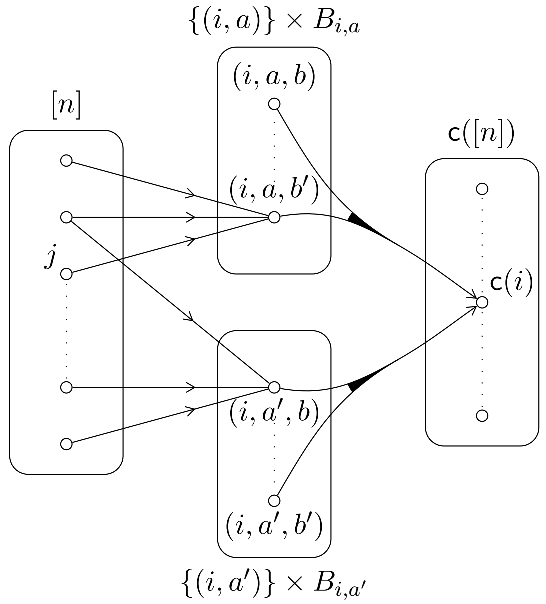

Figure 5.3: Hypergraph $\mathcal{G}_s^-$ associated with F^-

Lemma 5.19 that, for all subsets $I \subset [n]$, we have

$$\operatorname{reach}(I, \mathcal{G}^+) = \operatorname{reach}(I, \overline{\mathcal{G}}_s^+) \cap [n] \enspace ,$$
$$\operatorname{reach}(I, \mathcal{G}^-) = \operatorname{reach}(I, \overline{\mathcal{G}}_s^-) \cap [n] \enspace .$$

Thus, all the results of the previous sections, involving the hypergraphs $(\mathcal{G}^+, \mathcal{G}^-)$ and the families $(\mathcal{F}^+, \mathcal{F}^-)$ can be stated with the syntactic hypergraphs $(\overline{\mathcal{G}}_s^+, \overline{\mathcal{G}}_s^-)$. In particular we have the following characterization of dominions, which uses a stronger notion of invariance: in a directed hypergraph $\mathcal{G} = (N, E)$, if $N' \subset N$, we say that a subset I of N' is *invariant* in the hypergraph $\mathcal{G}$ *relatively* to N' if it contains all the nodes of N' that are reachable from itself, that is, $\operatorname{reach}(I, \mathcal{G}) \cap N' = I$.

Corollary 5.24. *Let Γ be a perfect-information finite stochastic game with state space $[n]$. Let $(\mathcal{F}^+, \mathcal{F}^-)$ and $(\overline{\mathcal{G}}_s^+, \overline{\mathcal{G}}_s^-)$ be the families of subsets of states and the syntactic hypergraphs, respectively, associated with Γ. Then, a nonempty subset $I \subset [n]$ is a dominion of player* MIN *(i.e., $I \in \mathcal{F}^+$) if, and only if, its complement in $[n]$ is an invariant set in the hypergraph $\overline{\mathcal{G}}_s^+$ relatively to $[n]$:*

$$\operatorname{reach}(I^c, \overline{\mathcal{G}}_s^+) \cap [n] = I^c \enspace .$$

Likewise, a nonempty subset $J \subset [n]$ is a dominion of player MAX *(i.e., $J \in \mathcal{F}^-$) if, and only if, its complement in $[n]$ is an invariant set in the hypergraph $\overline{\mathcal{G}}_s^-$ relatively to $[n]$:*

$$\operatorname{reach}(J^c, \overline{\mathcal{G}}_s^-) \cap [n] = J^c \enspace .$$

A characterization of ergodicity of stochastic games with perfect information follows.

Corollary 5.25. *A perfect-information finite stochastic game with state space $[n]$ is ergodic if, and only if, there do not exist nonempty disjoint subsets of states that are invariant in $\overline{\mathcal{G}}_s^+$ and $\overline{\mathcal{G}}_s^-$, respectively, relatively to $[n]$.*

Furthermore, the images of dominions by the Galois connection $(\Phi, \Phi^\star)$ can be effectively computed.

Corollary 5.26. *Let Γ be a perfect-information finite stochastic game with state space $[n]$. Let $(\mathcal{F}^+, \mathcal{F}^-)$ and $(\overline{\mathcal{G}}_s^+, \overline{\mathcal{G}}_s^-)$ be the families of subsets of states and the syntactic hypergraphs, respectively, associated with Γ. Then, for all $I \in \mathcal{F}^+$ (resp., $J \in \mathcal{F}^-$), $\Phi(I)$ (resp., $\Phi^\star(J)$) is given by the complement in $[n]$ of all the nodes of $[n]$ that are reachable from I (resp., J) in $\overline{\mathcal{G}}_s^-$ (resp., $\overline{\mathcal{G}}_s^+$):*

$$\Phi(I) = [n] \setminus \operatorname{reach}(I, \overline{\mathcal{G}}_s^-) \quad \Big(\textit{resp., } \Phi^\star(J) = [n] \setminus \operatorname{reach}(J, \overline{\mathcal{G}}_s^+)\Big) \enspace .$$

5.4.2 Checking ergodicity

According to Theorem 4.6, the negation of the following problem is equivalent to the next one.

Problem (NontrivialFP). Does a given payment-free Shapley operator F (5.8) associated with a perfect-information finite stochastic game have a nontrivial fixed point, that is, does there exist a vector $u \in \mathbb{R}^n \setminus \mathbb{R}e$ such that $u = F(u)$?

Problem (Ergodicity). Is a given perfect-information finite stochastic game Γ ergodic?

It is known that in a directed hypergraph $\mathcal{G}$, the set of reachable nodes from a set I can be computed in $O(\operatorname{size}(\mathcal{G}))$ time, see [GLNP93]. Hence, from Corollary 5.24 and Corollary 5.26 we get the following.

Proposition 5.27. *Let Γ be a perfect-information finite stochastic game with state space $[n]$, and let $(\mathcal{F}^+, \mathcal{F}^-)$ be the families of subsets of states associated with Γ. For any subsets $I, J \subset [n]$, checking that $I \in \mathcal{F}^-$ and $J \in \mathcal{F}^+$ can be done in $O(n^2|A|)$ and $O(n^2|A||B|)$ time, respectively. Moreover, $\Phi(I)$ and $\Phi^\star(J)$ can be evaluated in $O(n^2|A||B|)$ and $O(n^2|A|)$ time, respectively.* □

Using Proposition 5.27 and Corollary 5.8, we obtain the following result, which shows that checking the ergodicity of a game is fixed-parameter tractable: if the dimension is fixed, we can solve it in a time which is polynomial in the input size. Thus, for instances of moderate dimension but with large action spaces, ergodicity condition can be checked efficiently.

Theorem 5.28. *The ergodicity of a perfect-information finite stochastic game with state space $[n]$, that is the property "$\widehat{T}$ has only trivial fixed points", can be checked in $O(2^n n^2 |A||B|)$ time.* □

Problem **NontrivialFP** has already been addressed in the deterministic case with finite action spaces by Yang and Zhao [YZ04]. Suppose indeed that in the expression (5.8), the support of each transition probability is concentrated on just one state and consider the restriction of such an operator to the set of Boolean vectors $\{0, 1\}^n$. Then, we obtain a monotone Boolean operator.

Recall that a Boolean operator, defined on the set of Boolean vectors $\{0, 1\}^n$, is expressed using the logical operators `AND`, `OR` and `NOT`. Monotone Boolean operators are operators the expression of which involves only `AND` and `OR` operators. The latter can be interpreted as $\min$ and $\max$ operators, respectively. So, deterministic payment-free Shapley operators are equivalent to monotone Boolean operators and Problem **NontrivialFP** can be expressed in a simpler form.

Problem (MonBoolFP). Does a given monotone Boolean operator have a nontrivial fixed point, that is, different from the zero vector and the unit vector?

Theorem 5.29 ([YZ04, Cor. 1])**.** *Problem **MonBoolFP** is NP-complete.*

Using this result and the characterizations of the previous sections, we obtain the following.

Corollary 5.30. *Problem **NontrivialFP** is NP-complete.*

Proof. As a direct consequence of Theorem 5.29, we get that Problem **NontrivialFP** is NP-hard. We now show that it is in NP. Suppose that a payment-free Shapley operator F (5.8) has a nontrivial fixed point $u \in \mathbb{R}^n$. Then $\arg\min u$ and $\arg\max u$ are proper subsets of states, and by Lemma 5.5, we have $\arg\min u \in \mathcal{F}^+$ and $\Phi(\arg\min u) \supset \arg\max u \neq \emptyset$. Hence, $\arg\min u$ is a proper subset of states in $\mathcal{F}^+$ such that $\Phi(\arg\min u)$ is nonempty, and we know by Corollary 5.8 that these conditions are sufficient to guarantee the existence of a nontrivial fixed point. Furthermore, these conditions can be checked in polynomial time, as a consequence of Proposition 5.27. Hence, $\arg\min u$ is a short certificate to Problem **NontrivialFP**. □

5.4.3 Mean-payoff vectors with prescribed extrema

A way to analyze Problem **NontrivialFP** would be to characterize the set of fixed points $\mathcal{W} := \{w \in \mathbb{R}^n \mid F(w) = w\}$ of a payment-free Shapley operator F. This problem also arises in several other situations. First, in Proposition 4.4, we have shown that $\mathcal{W}$ is exactly the set of possible mean-payoff vectors of the parametric game $\Gamma(r) := ([n], A, B, K_A, K_B, r, p)$ when the payment function r varies. Next, $\mathcal{W}$ allows one to determine the set $\mathcal{E}$ of solutions u of the ergodic equation $T(u) = \lambda e + u$. Indeed, it is shown in [AGN16] that if the Shapley operator T is piecewise affine, if u is any point in $\mathcal{E}$ and if $\mathcal{V}$ is a neighborhood of 0, then

$$\mathcal{E} \cap (u + \mathcal{V}) = u + \{w \in \mathcal{V} \mid F(w) = w\} = u + (\mathcal{V} \cap \mathcal{W}) \ ,$$

where F is a payment-free Shapley operator (namely the semidifferential of T at point u). Hence, the local study of the ergodic equation reduces to the characterization of the fixed-point set $\mathcal{W}$.

Fixed points with prescribed argmin

In an attempt to understand the structure of the set of fixed points of a payment-free Shapley operator, we shall consider the following simpler problem.

Problem (MinFP). Let I be any subset of $[n]$. Does a given payment-free Shapley operator (5.8) associated with a perfect-information finite stochastic game have a fixed point u satisfying $\arg\min u = I$?

We know from Lemma 5.5 that a necessary condition is $I \in \mathcal{F}^+$. Under Assumption 5.B, this is equivalent to $F^+(e_{I^c}) \leqslant e_{I^c}$ (Lemma 5.18). We next show that there is a stronger necessary condition.

Lemma 5.31. *Let $F : \mathbb{R} \to \mathbb{R}^n$ be the payment-free Shapley operator associated with a stochastic game satisfying Assumption 5.B, and let $I \subset [n]$. Suppose that F has a fixed point u verifying $\arg\min u = I$. Then, $F^+(e_{I^c}) = e_{I^c}$.*

Proof. Let u be a fixed point of F verifying $\arg\min u = I$. If $I = [n]$, the conclusion of the lemma is trivial, so we may assume that $I \neq [n]$. Furthermore, we may suppose, without loss of generality, that $\min_{i \in [n]} u_i = 0$ and $\max_{i \in [n]} u_i = 1$, so that $u \leqslant e_{I^c}$. Since $F \leqslant F^+$, we get $u = F(u) \leqslant F^+(u) \leqslant F^+(e_{I^c})$. The last vector is Boolean, so this inequality implies $e_{I^c} \leqslant F^+(e_{I^c})$. Moreover, according to Lemma 5.5 and Lemma 5.18, we already know that $F^+(e_{I^c}) \leqslant e_{I^c}$. Hence the result. □

We continue with another necessary condition.

Lemma 5.32. *Let $F : \mathbb{R} \to \mathbb{R}^n$ be the payment-free Shapley operator associated with a stochastic game satisfying Assumption 5.B, and let $I \in \mathcal{F}^+$. If $\Phi(I) = \emptyset$, then F does not have any nontrivial fixed point u satisfying $I \subset \arg\min u$.*

Proof. Suppose on contrary that there is a nontrivial fixed point u such that $I \subset \arg\min u$. Let $I' := \arg\min u$ and $J := \arg\max u$. We know from Lemma 5.5 that $I' \in \mathcal{F}^+$, $J \in \mathcal{F}^-$ and that $J \subset \Phi(I')$. Since $I \subset I'$, we have $\Phi(I') \subset \Phi(I)$. Hence $J \subset \Phi(I)$, and since $J \neq \emptyset$, we get a contradiction. □

If $I = \emptyset$, the answer to Problem **MinFP** is trivially negative, and if $I = [n]$ it is trivially positive. Assume now that I is a proper subset of $[n]$. The above results show that a necessary condition to have a positive answer to problem **MinFP** is that $I \in \mathcal{F}^+$ and $\Phi(I) \neq \emptyset$. Moreover, by Lemma 5.6, a sufficient condition to have a positive answer to problem **MinFP** is that I is closed with respect to the Galois connection $(\Phi, \Phi^\star)$. It remains to examine the case in which $I \in \mathcal{F}^+$ is a proper subset, with $\Phi(I) \neq \emptyset$, and distinct from its closure $\overline{I} := \Phi^\star \circ \Phi(I)$ with respect to the Galois connection $(\Phi, \Phi^\star)$. Note that we must have, in particular, $\overline{I} \neq [n]$, otherwise we would have $\Phi(I) = \Phi(\overline{I}) = \emptyset$.

In the remainder of the subsection, F is the payment-free Shapley operator (5.8) associated with a perfect-information finite stochastic game Γ (hence, satisfying Assumption 5.B). We define a reduced operator $F^* : \mathbb{R}^{\overline{I}} \to \mathbb{R}^{\overline{I}}$ as follows. According to the game-theoretic interpretation (Proposition 5.15), we know that MIN can force the state of the game Γ to stay in $\overline{I}$ from any initial position in $\overline{I}$. Hence, we consider the actions of MIN that achieve this goal: for every $i \in \overline{I}$, let

$$A_i^* := \left\{a \in A_i \mid \forall b \in B_{i,a},\ \ \forall j \notin \overline{I},\ \ p(j \mid i, a, b) = 0\right\} \ .$$

These sets are nonempty since $\overline{I} \in \mathcal{F}^+$. Alternatively, A_i^* can be defined as the set of actions $a \in A_i$ which attain the minimum in the minimax formula (5.8) of $F_i\big(e_{[n]\setminus\overline{I}}\big)$, that is,

$$A_i^* = \left\{a \in A_i \;\middle|\; \max_{b \in B_{i,a}} \sum_{j \notin \overline{I}} p(j \mid i, a, b) = F_i\big(e_{[n]\setminus\overline{I}}\big) = 0\right\} \ .$$

Given $x \in \mathbb{R}^n$ and $K \subset [n]$, let us denote by x_K the restriction of x to $\mathbb{R}^K$. Then, we define the reduced map F^* from $\mathbb{R}^{\overline{I}}$ to itself, whose ith component ($i \in \overline{I}$) is given by

$$F_i^*(x) := \min_{a \in A_i^*} \max_{b \in B_{i,a}} \sum_{j \in \overline{I}} x_j \, p(j \mid i, a, b) \ , \quad x \in \mathbb{R}^{\overline{I}} \ .$$

The latter map is a payment-free Shapley operator on $\mathbb{R}^{\overline{I}}$. Indeed, from the definition of A_i^*, any transition probability $p(\cdot \mid i, a, b)$ is supported by $\overline{I}$ as soon as $a \in A_i^*$ with $i \in \overline{I}$. In particular, we have, for all $i \in \overline{I}$,

$$F_i^*(x_{\overline{I}}) = \min_{a \in A_i^*} \max_{b \in B_{i,a}} \sum_{j \in [n]} x_j \, p(j \mid i, a, b) \ , \quad \forall x \in \mathbb{R}^n \ . \tag{5.9}$$

Theorem 5.33. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be the payment-free Shapley operator associated with a perfect-information finite stochastic game. Let $I \in \mathcal{F}^+$ be proper, such that $\Phi(I) \neq \emptyset$ and $I \neq \overline{I}$. Then F has a fixed point the $\arg\min$ of which is I if, and only if, the same holds for the reduced operator F^*.*

Proof. We first show the "only if" part of the theorem. Let u be a fixed point of F such that $I = \arg\min u$. Recall that $I \neq [n]$ by hypothesis. So we may suppose that $\max_{i\in[n]} u_i = 1$ and $\min_{i\in[u]} u_i = 0$.

It follows from (5.9) that $u_i = F_i(u) \leqslant F_i^*(u_{\overline{I}})$ for all $i \in \overline{I}$, so that $(F^*)^\omega(u_{\overline{I}})$ exists. Let us denote this limit point by v. It is a fixed point of F^* and it satisfies $u_{\overline{I}} \leqslant v$. As a consequence, $v_i > 0$ for every $i \in \overline{I} \setminus I$.

Furthermore, Lemma 5.5 implies that $I \in \mathcal{F}^+$, meaning that $F(e_{[n]\setminus I}) \leqslant e_{[n]\setminus I}$. Then, for all $i \in I$, there exists some action $a \in A_i$ such that $p([n] \setminus I \mid i, a, b) = 0$ for all actions $b \in B_{i,a}$. Since $I \subset \overline{I}$, this implies that $p(j \mid i, a, b) = 0$ for all $j \in [n] \setminus \overline{I}$, and since this holds for all $b \in B_{i,a}$, we deduce that $a \in A_i^*$. Hence,

$$\min_{a\in A_i^*} \max_{b\in B_{i,a}} \sum_{j\notin I} p(j \mid i, a, b) = 0 \ ,$$

which yields $F_i^*(e_{\overline{I}\setminus I}) = 0$ for all $i \in I$. Therefore $F^*(e_{\overline{I}\setminus I}) \leqslant e_{\overline{I}\setminus I}$, and since $u_{\overline{I}} \leqslant e_{\overline{I}\setminus I}$, it follows that

$$v = (F^*)^\omega(u_{\overline{I}}) \leqslant (F^*)^\omega(e_{\overline{I}\setminus I}) \leqslant e_{\overline{I}\setminus I} \ .$$

Hence $v_i = 0$ for every $i \in I$, which shows that $\arg\min v = I$.

We now prove the "if" part of the theorem. Assume that F^* has a fixed point v such that $\arg\min v = I$. We may suppose that $\max_{i\in[n]} v_i = 1$ and $\min_{i\in[n]} v_i = 0$. Let $w := F^\omega(e_{[n]\setminus\overline{I}})$. We know from Lemma 5.6 that w is a fixed point of F verifying $\arg\min w = \overline{I}$. Thus, it satisfies $w_{\overline{I}} = 0$ and $w_j > 0$ for every $j \in [n] \setminus \overline{I}$, hence there exists some constant $\alpha > 0$ such that $w \geqslant \alpha e_{[n]\setminus\overline{I}}$.

We next use the notions of semidifferentiability and semiderivative, referring the reader to [RW98, AGN16] for the definition of these notions and for their basic properties. Since the action spaces are finite, F is piecewise affine and so, it is semidifferentiable at point w. Furthermore, denoting by F'_w its semiderivative at w, there is a neighborhood $\mathcal{V}$ of 0 such that

$$F(w + x) = F(w) + F'_w(x) \ , \quad \forall x \in \mathcal{V} \ . \tag{5.10}$$

We next give a formula for F'_w. For every $i \in [n]$, let

$$A_i(w) := \left\{ a \in A_i \;\middle|\; \max_{b\in B_{i,a}} \sum_{j\in[n]} w_j \, p(j \mid i, a, b) = F_i(w) \right\}$$

and for all $a \in A_i(w)$, let

$$B_{i,a}(w) := \left\{ b \in B_{i,a} \;\middle|\; \sum_{j\in[n]} w_j \, p(j \mid i, a, b) = F_i(w) \right\} \ .$$

Then, the ith component of F'_w is given by

$$\left[F'_w(x)\right]_i = \min_{a\in A_i(w)} \max_{b\in B_{i,a}(w)} \sum_{j\in[n]} x_j \, p(j \mid i, a, b) \ , \quad x \in \mathbb{R}^n \ .$$

Observe that for each $i \in \overline{I}$, we have $A_i(w) = A_i^*$ and $B_{i,a}(w) = B_i$ for all $a \in A_i(w)$. This is because, for all $i \in \overline{I}$ we have $F_i(w) = w_i = 0$ and $w_j > 0$ for all $j \notin \overline{I}$, so that $a \in A_i(w)$ if, and only if, $\max_{b\in B_{i,a}} p([n] \setminus \overline{I} \mid i, a, b) = 0$ and $b \in B_{i,a}(w)$ if, and only if, $p([n] \setminus \overline{I} \mid i, a, b) = 0$. Then, using (5.9), we obtain that

$$[F'_w(x)]_{\overline{I}} = F^*(x_{\overline{I}}) \ , \quad \forall x \in \mathbb{R}^n \ . \tag{5.11}$$

We now introduce the vector $z \in [0,1]^n$ given by $z_{\overline{I}} = v$ and $z_{[n]\setminus\overline{I}} = 0$. By the above property (5.11) of F'_w, we get that $[F'_w(z)]_{\overline{I}} = F^*(v) = v = z_{\overline{I}}$. Moreover, since F'_w is a payment-free operator and $z \geqslant 0$, we get that $F'_w(z) \geqslant 0$, so $F'_w(z) \geqslant z$. Hence, the limit point $\bar{z} := (F'_w)^\omega(z)$ exists and is a fixed point of F'_w, belonging to $[0,1]^n$. Again, by the above property (5.11) of F'_w, we get that $[(F'_w)^k(z)]_{\overline{I}} = F^*([(F'_w)^{k-1}(z)]_{\overline{I}})$ for all integers $k \geqslant 1$, so that by induction $[(F'_w)^k(z)]_{\overline{I}} = v$, and finally $\bar{z}_i = v_i$ for all $i \in \overline{I}$.

Choose $\varepsilon > 0$ small enough so that $\varepsilon\bar{z}$ is in $\mathcal{V}$ and let $u = w + \varepsilon\,\bar{z}$. Then, from (5.10), we get that $F(u) = F(w) + \varepsilon F'_w(\bar{z}) = w + \varepsilon\bar{z} = u$, where we used the fact that F'_w is positively homogeneous. Then u is a fixed point of F. Moreover, by construction $u = w + \varepsilon\,\bar{z} \geqslant w$ and $u \geqslant \varepsilon\bar{z}$, and since $\arg\min w = \overline{I}$ and $\arg\min \bar{z} \cap \overline{I} = I$, we deduce that $u_I = 0$ and $u_j > 0$ for every $j \in [n] \setminus I$, that is, $\arg\min u = I$. □

The previous result, together with the observations made before it, lead to Algorithm 1 below, which solves Problem **MinFP**, as detailed in Theorem 5.34. There, if Γ denotes a finite stochastic game, we write Γ^* the reduced game with, in particular, state space $\overline{I}$ and set of admissible actions A_i^* for player MIN when in state $i \in \overline{I}$. Also, if F is a payment-free Shapley operator, we write $(\Phi_F, \Phi_F^\star)$ the Galois connection associated with that operator.

Algorithm 1:

input : perfect-information finite stochastic game $\Gamma := (S, A, B, K_A, K_B, 0, p)$,
corresponding payment-free Shapley operator $F : \mathbb{R}^S \to \mathbb{R}^S$, subset $I \subset S$.
output: answer to Problem **MinFP**.

if $I = \emptyset$ **then**
  return false
else if $I = S$ **then**
  return true
else
  repeat
    if $\big(F^+(e_{I^c}) \neq e_{I^c}$ **or** $\Phi_F(I) = \emptyset \big)$ **then**
      return false
    else if $\Phi_F^\star \circ \Phi_F(I) = I$ **then**
      return true
    else
      $S \leftarrow \Phi_F^\star \circ \Phi_F(I), \Gamma \leftarrow \Gamma^*, F \leftarrow F^*$
    end
  end
end

Theorem 5.34. *Algorithm 1 solves Problem **MinFP** in $O(n^3|A||B|)$ time.*

Proof. The fact that Algorithm 1 provides the right answer to Problem **MinFP** is a direct consequence of Lemma 5.31, Lemma 5.32, Lemma 5.6 and Theorem 5.33.

We next show that it stops after at most n iterations of the loop. Suppose that during the execution of a loop, the first two conditions (which are stopping criteria) are not satisfied. Then the closure of I with respect to the Galois connection $(\Phi_F, \Phi_F^\star)$ associated with F is a proper subset of states. Hence, the cardinality of the state space for the reduced operator F^* is strictly less than the one for F. Moreover, each operation in the loop requires at most $O(n^2|A||B|)$ time (see Proposition 5.27). □

Fixed points with prescribed argmin and argmax

So far, we have only considered the problem with a single constraint on the fixed point, concerning the indices of the minimal entries. The dual problem, concerning the maximal entries of fixed points, is equivalent. We address now a mixed-condition problem.

Problem (MinMaxFP). Let I and J be nonempty disjoint subsets of $[n]$. Does a given payment-free Shapley operator (5.8) associated with a perfect-information finite stochastic game have a fixed point u satisfying $\arg\min u = I$ and $\arg\max u = J$?

The following theorem shows that Problem **MinMaxFP** can be split in one instance of Problem **MinFP** and one instance of its dual problem.

Theorem 5.35. *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be the payment-free Shapley operator associated with a perfect-information finite stochastic game. Let $I \in \mathcal{F}^+$ and $J \in \mathcal{F}^-$ be two nonempty disjoint subsets. Then F has a fixed point u satisfying $\arg\min u = I$ and $\arg\max u = J$ if, and only if, F has fixed points v and w satisfying $\arg\min v = I$ and $\arg\max w = J$, respectively.*

Proof. We only need to prove the "if" part of the theorem. Suppose that F has fixed points v and w such that $\arg\min v = I$ and $\arg\max w = J$, respectively. Then, we may impose

$$\min_{i\in[n]} v_i = 0 \ , \quad \max_{i\in[n]} v_i = \min_{i\in[n]} w_i = 1/2 \ , \quad \max_{i\in[n]} w_i = 1 \ .$$

Let $\mathcal{L} := \{z \in \mathbb{R}^n \mid v \vee e_J \leqslant z \leqslant w \wedge e_{I^c}\}$, where $\vee$ and $\wedge$ stands for the supremum and the infimum operators, respectively. Put in words, $\mathcal{L}$ is the set of all elements in the hypercube $[0,1]^n$ whose entries are 0 on I, 1 on J and between the ones of v and w elsewhere. In particular, the entries outside I or J of the vectors in $\mathcal{L}$ are in $(0,1)$.

The set $\mathcal{L}$ is a complete lattice. We next show that it is invariant by F. Since $J \in \mathcal{F}^-$, we have

$$v \vee e_J \leqslant F(v) \vee F(e_J) \leqslant F(v \vee e_J) \ ,$$

and since $I \in \mathcal{F}^+$, we have

$$w \wedge e_{I^c} \geqslant F(w) \wedge F(e_{I^c}) \geqslant F(w \wedge e_{I^c}) \ .$$

Hence, $v \vee e_J \leqslant z \leqslant w \wedge e_{I^c}$ implies

$$v \vee e_J \leqslant F(v \vee e_J) \leqslant F(z) \leqslant F(w \wedge e_{I^c}) \leqslant w \wedge e_{I^c} \ ,$$

which shows that $\mathcal{L}$ is invariant by F. Thus, since F is order-preserving, Tarski's fixed-point theorem guarantees the existence of a fixed point of F in $\mathcal{L}$. □

Corollary 5.36. *Problem **MinMaxFP** can be solved in $O(n^3|A||B|)$ time.*

Proof. According to Theorem 5.35, Problem **MinMaxFP** can be solved by two instances of Problem **MinFP**, one with inputs F and I, one with inputs $\widetilde{F}$ (Remark 5.3) and J. □

Summary of complexity results

The following table summarizes the results of this section.

Problem	Complexity class	
MonBoolFP	NP-complete	([YZ04, Cor. 1])
NontrivialFP	NP-complete	(Cor. 5.30)
MinFP	P	(Th. 5.34)
MinMaxFP	P	(Cor. 5.36)

Table 5.1: Complexity class of fixed-point problems

5.5 Example

5.5.1 Checking ergodicity

Let us consider the zero-sum stochastic game with perfect information defined by the graph represented in Figure 5.4. There are four states represented by gray nodes. A token is initially placed in one of these nodes. At each stage, the token is moved along the edges of the graph until it reaches another state, according to the following rule: player MIN moves the token at circle nodes, player MAX at square ones and at diamond nodes, an edge is selected at random according to the probabilities indicated on the edges starting from the node. A payment only occurs for the edges starting from a MAX node (its value is given by the label attached to them).

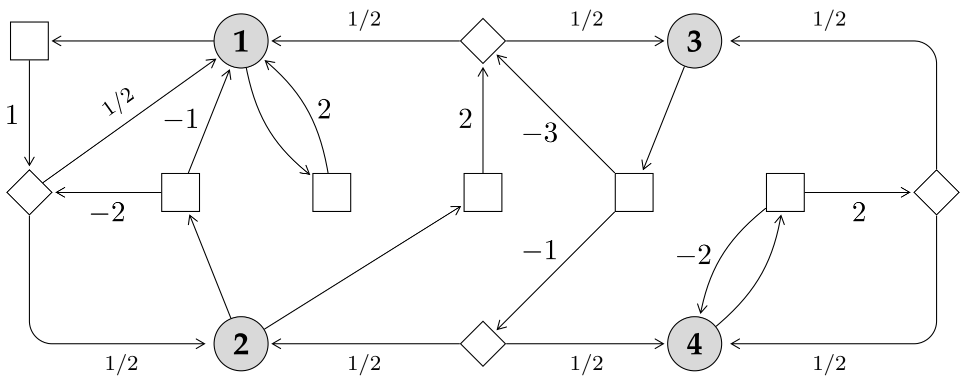

Figure 5.4: Graph on which is played a perfect-information stochastic game

The Shapley operator of this game is

$$T(x) = \begin{pmatrix} (2+x_1) \wedge (1+\frac{1}{2}(x_1+x_2)) \\ ((-2+\frac{1}{2}(x_1+x_2)) \vee (-1+x_1)) \wedge (2+\frac{1}{2}(x_1+x_3)) \\ (-3+\frac{1}{2}(x_1+x_3)) \vee (-1+\frac{1}{2}(x_2+x_4)) \\ (-2+x_4) \vee (2+\frac{1}{2}(x_3+x_4)) \end{pmatrix} , \quad x \in \mathbb{R}^4 .$$

It can be checked that the ergodic equation (2.5) is solvable, with ergodic constant $\lambda = 1/3$ and bias vector $u = (4/3, 0, 2/3, 4)^\intercal$. Let us decide whether this game is ergodic.

This can be easily done following the game-theoretic interpretation of Section 5.3. Indeed, observe on Figure 5.4 that the dominions of MIN are $\{1\}$, $\{1,2\}$ and $\{1,2,3,4\}$, since, apart from the whole set of states, MIN can always make sure that the state remains in $\{1\}$ or in $\{1,2\}$. As for MAX, its dominions are $\{4\}$ and $\{1,2,3,4\}$. Thus, there exists at least one pair of disjoint dominions of MIN and MAX, e.g., $\{1\}$ and $\{4\}$, and so we conclude from Corollary 5.16 that the game is not ergodic.

Alternatively, and in order to illustrate the Galois connection of Section 5.1, we can study the fixed points of the recession operator of T, denoted here by F and given by

$$F(x) = \begin{pmatrix} x_1 \wedge \frac{1}{2}(x_1 + x_2) \\ \left(\frac{1}{2}(x_1 + x_2) \vee x_1\right) \wedge \frac{1}{2}(x_1 + x_3) \\ \frac{1}{2}(x_1 + x_3) \vee \frac{1}{2}(x_2 + x_4) \\ x_4 \vee \frac{1}{2}(x_3 + x_4) \end{pmatrix} , \quad x \in \mathbb{R}^4 \tag{5.12}$$

From Proposition 5.15 we readily get that the families $\mathcal{F}^+$ and $\mathcal{F}^-$ associated with F are

$$\mathcal{F}^+ = \big\{\emptyset, \{1\}, \{1,2\}, \{1,2,3,4\}\big\} \quad \text{and} \quad \mathcal{F}^- = \big\{\emptyset, \{4\}, \{1,2,3,4\}\big\} \ .$$

We can also obtain a description of the latter families by constructing the Boolean abstractions of F:

$$F^+(x) = \begin{pmatrix} x_1 \\ x_1 \vee (x_2 \wedge x_3) \\ x_1 \vee x_3 \vee x_2 \vee x_4 \\ x_3 \vee x_4 \end{pmatrix} \quad \text{and} \quad F^-(x) = \begin{pmatrix} x_1 \wedge x_2 \\ x_1 \wedge x_3 \\ (x_1 \wedge x_3) \vee (x_2 \wedge x_4) \\ x_4 \end{pmatrix} ,$$

and check that

$$F^+(e_{\{2,3,4\}}) \leqslant e_{\{2,3,4\}} \ , \quad F^+(e_{\{3,4\}}) \leqslant e_{\{3,4\}} \quad \text{and} \quad F^-(e_{\{4\}}) \geqslant e_{\{4\}} \ .$$

Then, by definition of the Galois connection, we have

$$\Phi(\{1\}) = \Phi(\{1,2\}) = \{4\} \quad \text{and} \quad \Phi^\star(\{4\}) = \{1,2\} \ .$$

So, $(\{1,2\},\{4\})$ is a pair of conjugate subsets with respect to the Galois connection, which proves (Theorem 5.7 together with Theorem 4.6) that the game is not ergodic.

5.5.2 Finding a fixed point with prescribed argmin

We now address the problem of finding fixed points of F with fixed $\arg\min$. Since $e_{\{2,3,4\}}$ and $e_{\{3,4\}}$ are the only nontrivial fixed points of F^+, we know from Lemma 5.31 that $\{1\}$ and $\{1,2\}$ are the only possible candidates for nontrivial $\arg\min$.

The set $\{1,2\}$ is closed with respect to the Galois connection $(\Phi,\Phi^\star)$. Thus, according to Lemma 5.6, F has a fixed point whose $\arg\min$ is $\{1,2\}$. Moreover, its $\arg\max$ can only be $\{4\}$. Check that the vector $(0,0,1/2,1)^\intercal$ is a fixed point with these properties.

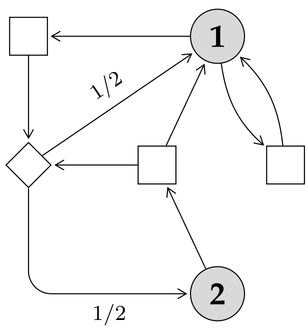

Figure 5.5: Reduced game associated with F^*

As for the set $\{1\}$, we cannot conclude directly, since the hypotheses of Lemma 5.32 or Lemma 5.6 are not satisfied. Hence, following Theorem 5.33, we need to construct the reduced operator F^* defined on $\mathbb{R}^{\{1,2\}}$ ($\{1,2\}$ being the closure of $\{1\}$) by

$$F^*(x) = \begin{pmatrix} x_1 \wedge \frac{1}{2}(x_1 + x_2) \\ x_1 \vee \frac{1}{2}(x_1 + x_2) \end{pmatrix} \quad , \quad x \in \mathbb{R}^2$$

The directed graph associated with this operator is represented in Figure 5.5.

We check that for this reduced operator we have

$$\mathcal{F}^+ = \big\{\emptyset, \{1\}, \{1,2\}\big\} \quad \text{and} \quad \mathcal{F}^- = \big\{\emptyset, \{1,2\}\big\} \ .$$

Hence, $\Phi(\{1\}) = \emptyset$ and by Lemma 5.32, we know that F^* has no fixed point the $\arg\min$ of which is $\{1\}$. According to Theorem 5.33, the same holds for F.

We conclude that any nontrivial fixed point u of F verifies $u_1 = u_2 < u_3 < u_4$, and so, from (5.12), also verifies $u_3 = \frac{1}{2}(u_2 + u_4)$. Finally, note that if the value of the payments are considered as parameters, then, according to Proposition 4.3 and Proposition 4.4, all the realizable mean payoff vectors χ are characterized by

$$\chi_1 = \chi_2 \leqslant \chi_4, \quad \chi_3 = \frac{1}{2}(\chi_1 + \chi_4) \ .$$

Chapter 6

A game-abstraction condition for ergodicity of stochastic games

In this chapter, we provide a combinatorial condition for the boundedness of all slice spaces in Hilbert's seminorm for any Shapley operator T on $\mathbb{R}^n$. This condition is based on a simplified version of a stochastic game (a game abstraction), constructed from T.

The results presented in this chapter have been partly announced in the CDC conference proceedings [AGH15b], with a different point of view entirely based on graph-theoretic aspects.

6.1 Game abstraction, dominions and slice spaces

6.1.1 Dominion condition and boundedness of slice spaces

Game abstraction

Let us fix a Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$. We associate to T an auxiliary two-player stochastic game, denoted by Γ_T, defined in the following way. The state space is $[n]$ and when in state i, player Min can choose a subset of states $I \subset [n]$ such that

$$\lim_{\alpha \to +\infty} T_i(\alpha e_{I^c}) < +\infty \enspace .$$

Likewise, in state i, player Max can choose a subset of states $J \subset [n]$ such that

$$\lim_{\alpha \to -\infty} T_i(\alpha e_{J^c}) > -\infty \enspace .$$

At each stage, Min and Max choose simultaneously a subset I and J, respectively, and the state at the next stage is drawn uniformly in $I \cap J$. Note that here, we are only interested in the dynamic of the state, and therefore we do not need to define a payoff function.

Let us make some observations. First, since T is monotone, the condition for I to be in the action space of Min in state i is equivalent to the nondecreasing sequence $(T_i(ke_{I^c}))_{k\in\mathbb{N}}$ being bounded above. Likewise, J is in the action space of Max in state i if, and only if, the nonincreasing sequence $(T_i(-ke_{J^c}))_{k\in\mathbb{N}}$ is bounded below. The monotonicity of T also implies that if any player can choose a subset I in a given state, then he can also choose any other subset $K \supset I$.

Second, the action spaces of the two players are never empty in any state, since they contain the set of all states $[n]$. Also observe that no action space contains the empty set.

Finally, suppose that in state i the subsets I and J chosen by MIN and MAX, respectively, are disjoint. Then, we have $e - e_{J^c} = e_J \leqslant e_{I^c}$. By monotonicity and additive homogeneity of T, we deduce that $\alpha + T_i(-\alpha e_{J^c}) \leqslant T_i(\alpha e_{I^c})$ for all scalars $\alpha \geqslant 0$. The latter inequality yields that $\lim_{\alpha\to+\infty} T_i(\alpha e_{I^c})$ and $\lim_{\alpha\to+\infty} T_i(-\alpha e_{J^c})$ cannot be both finite, a contradiction. Hence, we necessarily have $I \cap J \neq \emptyset$, which shows, along with the previous observation, that the game is well defined.

Dominions

In the game Γ_T, we call *dominion* a nonempty subset of states $D \subset [n]$ such that from any position in D, one player can force the state at the next stage to remain in D almost surely, whatever action the other player chooses. Equivalently, D is a dominion of one player in Γ_T if for all states $i \in D$, that player can pick a subset $I \subset D$. This is also equivalent to saying that in each state $i \in D$, the action space of the player contains the subset D.

The main result of this chapter characterizes the boundedness of slice spaces in Hilbert's seminorm in terms of these dominions. To that purpose, we shall need the following condition.

Assumption 6.C (Dominion condition). In Γ_T, players MIN and MAX have each a dominion, I and J respectively, such that $I \cap J = \emptyset$.

Remark 6.1. The notion of dominion defined here is the same as the one introduced in Section 5.3. The only difference lies in the game with respect to which it is considered. In this chapter, we do not assume that T has an explicit minimax representation, and we construct a game abstraction Γ_T, whereas in Section 5.3, the Shapley operator T arises as the dynamic programming operator of a given zero-sum stochastic game Γ, the data of which are known. However, when T arises explicitly from a zero-sum stochastic game Γ satisfying Assumption 5.B, then the dominions are the same in Γ_T and in Γ. Indeed, when in state i, player MIN can choose subset I in Γ_T if in the one-shot game associated with Γ, MIN is not "penalized" by an arbitrary large final payment of α in states outside I. In particular, if Assumption 5.B holds, it is easy to see that MIN can choose subset I in state i if, and only if, $\widehat{T}_i(e_{I^c}) = 0$. Thus, the subset I is a dominion of MIN in Γ_T if, and only if, $\widehat{T}(e_{I^c}) \leqslant e_{I^c}$, i.e., I is a dominion of MIN in Γ (see Proposition 5.15). A dual statement holds for player MAX.

We now state the main result.

Theorem 6.2. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a Shapley operator. All slice spaces $\mathcal{S}_\alpha^\beta(T)$ are bounded in Hilbert's seminorm if, and only if, the dominion condition 6.C does not hold in the game Γ_T.*

According to Theorem 4.21, the above result yields a combinatorial condition of ergodicity.

Corollary 6.3. *A Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$ is ergodic if, and only if, the dominion condition 6.C does not hold in the game Γ_T.*

6.1.2 Proof of Theorem 6.2

"if" part

We prove the converse statement. Suppose that there is some slice space $\mathcal{S}_\alpha^\beta(T)$ unbounded in Hilbert's seminorm. So there exists a sequence $(u^k)_{k\in\mathbb{N}}$ of vectors in $\mathcal{S}_\alpha^\beta(T)$ such that $\|u^k\|_{\mathrm{H}} \to$

$+\infty$ as $k \to +\infty$. We may assume, without loss of generality, that $\min_{i\in[n]} u_i^k = 0$ for all integers k, and we let

$$\tau^k := \max_{i\in[n]} u_i^k = \|u^k\|_{\mathrm{H}} \enspace .$$

Among all pairs of indices, there is at least one, (i_0, j_0), such that $u_{i_0}^k = 0$ and $u_{j_0}^k = \tau^k$ for infinitely many integers k. Then, up to the extraction of a subsequence, we may assume that the latter equalities hold for all integers.

We next explain how to construct a partition of $[n]$ in three subsets I, J, L such that there exists a strictly increasing function $\phi : \mathbb{N} \to \mathbb{N}$ for which we have:

– for all $i \in I$, the sequence $(u_i^{\phi(k)})_k$ is bounded;
– for all $j \in J$, the sequence $(\tau^{\phi(k)} - u_j^{\phi(k)})_k$ is bounded;
– for all $\ell \in L$, the sequences $(u_\ell^{\phi(k)})_k$ and $(\tau^{\phi(k)} - u_\ell^{\phi(k)})_k$ tend to ∞.

We proceed by induction, starting with $I_0 = \{i_0\}$, $J_0 = \{j_0\}$, $L_0 = \emptyset$, and letting ϕ_0 be the identity function. Then, assume that we have constructed three disjoint subsets, I_m, J_m and L_m for which the above statements are satisfied for some function $\phi_m : \mathbb{N} \to \mathbb{N}$, and such that $I_m \cup J_m \cup L_m \neq [n]$. Let $i \in [n] \setminus (I_m \cup J_m \cup L_m)$. If $\liminf_{k\to\infty} u_i^{\phi_m(k)}$ is finite, then there exists a bounded subsequence of $(u_i^{\phi_m(k)})$. In this case, we let $I_{m+1} = I_m \cup \{i\}$, $J_{m+1} = J_m$, and $L_{m+1} = L_m$. If $(u_i^{\phi_m(k)})$ tends to ∞ but $\liminf_{k\to\infty} \tau^{\phi_m(k)} - u_i^{\phi_m(k)}$ is finite, then there exists a bounded subsequence of $(\tau^{\phi_m(k)} - u_i^{\phi_m(k)})$. In that case, we let $J_{m+1} = J_m \cup \{i\}$, $I_{m+1} = I_m$ and $L_{m+1} = L_m$. In both cases, we denote by $\phi_{m+1} : \mathbb{N} \to \mathbb{N}$ the strictly increasing function such that $(u_i^{\phi_{m+1}(k)})$ is the aforementioned subsequence. If neither cases hold, then we let $L_{m+1} = L_m \cup \{i\}$, $I_{m+1} = I_m$, $J_{m+1} = J_m$ and $\phi_{m+1} = \phi_m$. The induction is finite and we obtain at the last step a partition of $[n]$ with the required properties. Furthermore, we know that I and J are nonempty.

For the sake of simplicity, assume now that ϕ is the identity function. For all integers k, let us denote

$$\rho^k := \min_{j\notin I} u_j^k \quad \text{and} \quad \sigma^k := \max_{i\notin J} u_i^k \enspace .$$

By construction, we have $\lim_{k\to\infty} \rho^k = \lim_{k\to\infty} \tau^k - \sigma^k = +\infty$. Let $M > 0$ be a joint upper bound of the sequences $(u_i^k)_k$ for all $i \in I$, and $(\tau^k - u_j^k)_k$ for all $j \in J$. Then we have, for all indices $i \in I$,

$$T_i(\rho^k e_{I^c}) \leqslant T_i(u^k) \leqslant \beta + u_i^k \leqslant \beta + M \enspace .$$

This proves that in the game Γ_T, MIN can choose subset I in each state $i \in I$. Hence, I is a dominion of MIN. Likewise, for all indices $j \in J$, we have

$$T_j\Big((\sigma^k - \tau^k)e_{J^c}\Big) \geqslant T_j(u^k - \tau^k e) \geqslant \alpha + u_j^k - \tau^k \geqslant \alpha - M \enspace ,$$

which proves that J is a dominion of MAX, and thus that the dominion condition holds in the game Γ_T.

"only if" part

We also prove the converse statement, and so we assume that the dominion condition holds in Γ_T. Let $I, J \subset [n]$ be disjoint dominions of MIN and MAX, respectively. Then, by definition of the game Γ_T, there exists a scalar $M > 0$ such that, for all $k \in \mathbb{N}$,

$$\begin{cases} T_i(0) \leqslant T_i(ke_{I^c}) \leqslant M \enspace , & \forall i \in I \enspace , \\ -M \leqslant T_j(-ke_{J^c}) \leqslant T_j(0) \enspace , & \forall j \in J \enspace . \end{cases} \tag{6.1}$$

We can assume that $M \geqslant \|T(0)\|_\infty$.

Let $L = [n] \setminus (I \cup J)$ and, for all integers k, let u^k be a vector of $\mathbb{R}^n$ such that

$$\begin{cases} u_j^k = k \ , & \forall j \in J \ , \\ -k \leqslant u_\ell^k \leqslant k \ , & \forall \ell \in L \ , \\ u_i^k = -k \ , & \forall i \in I \ . \end{cases} \tag{6.2}$$

Observe that $\|u^k\|_{\mathrm{H}} = 2k$. Moreover, by construction, we have $0 \leqslant u^k + ke \leqslant 2ke_{I^c}$ and $-2ke_{J^c} \leqslant u^k - ke \leqslant 0$. From the inequalities (6.1), we then deduce that

$$\begin{aligned} -M \leqslant T_i(0) \leqslant T_i(u^k) + k = T_i(u^k) - u_i^k \leqslant M \ , \quad \forall i \in I \ , \\ -M \leqslant T_j(u^k) - k = T_j(u^k) - u_j^k \leqslant T_j(0) \leqslant M \ , \quad \forall j \in J \ . \end{aligned}$$

If $L = \emptyset$, we readily have that $u^k \in \mathcal{S}_{-M}^M(T)$ for all $k \in \mathbb{N}$, so that $\mathcal{S}_{-M}^M(T)$ contains a sequence which is unbounded in Hilbert's seminorm, and the proof of the "only if" part is complete in this case.

Assume now that $L \neq \emptyset$. We next show that we can find scalars α, β for which it is possible to construct, for all $k \in \mathbb{N}$, a vector u^k satisfying (6.2) and such that $\alpha \leqslant T_\ell(u^k) - u_\ell^k \leqslant \beta$ for every $\ell \in L$. To that purpose, we shall use the Poincaré-Miranda theorem, a generalization of the intermediate value theorem, see [Kul97].

Theorem 6.4 (Poincaré-Miranda). *Let $f : [0,1]^n \to \mathbb{R}^n$ be a continuous map such that for each $i \in [n]$, $f_i(x) \leqslant 0$ whenever $x_i = 0$ and $f_i(x) \geqslant 0$ whenever $x_i = 1$. Then there exists a point $x^* \in [0,1]^n$ such that $f(x^*) = 0$.*

Let $k \in \mathbb{N}$ be a fixed integer and $G : [0,1]^L \to [-k,k]^n$ be the map given by

$$\begin{cases} G_j(x) = k \ , & \forall j \in J \ , \\ G_\ell(x) = -k + 2kx_\ell \ , & \forall \ell \in L \ , \\ G_i(x) = -k \ , & \forall i \in I \ . \end{cases}$$

Let $F : [0,1]^L \to \mathbb{R}^L$ be the map whose ℓth entry is defined by

$$F_\ell(x) = G_\ell(x) - T_\ell \circ G(x) + T_\ell(0) \ .$$

Since we have $-k + T_\ell(0) \leqslant T_\ell \circ G(x) \leqslant k + T_\ell(0)$ for every $x \in [0,1]^L$ and every $\ell \in L$, then we deduce that $F_\ell(x) \leqslant 0$ for all x such that $x_\ell = 0$. Likewise, $F_\ell(x) \geqslant 0$ for all x such that $x_\ell = 1$. Thus, according to the Poincaré-Miranda theorem, there exists a point $x^k \in [0,1]^L$ such that $F(x^k) = 0$. Fixing $u^k = G(x^k)$, we get that $T_\ell(u^k) - u_\ell^k = T_\ell(0)$ for all $\ell \in L$.

Hence, it is possible to choose a vector u^k satisfying (6.2), and such that $u^k \in \mathcal{S}_{-M}^M(T)$ for all integers $k \in \mathbb{N}$. So, $\mathcal{S}_{-M}^M(T)$ contains a sequence which is unbounded in Hilbert's seminorm, which conclude the proof of the "only if" part and of Theorem 6.2. ☐

6.1.3 Convex case

In this subsection, we consider a convex Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$, meaning that every coordinate function of T is convex. In this particular case, we show that Theorem 6.2 simplifies. More precisely, we next show that the two-player game Γ_T can be reduced to a one-player game in which the actions of MIN are not taken into account while the actions of MAX are essentially the same. However, since many actions of MAX in Γ_T are superfluous (recall that if he can choose a subset J, then he can also choose any subset $K \supset J$), we shall need to restrict his action spaces.

Before introducing this one-player game, let us give the following definition. For any index $i \in [n]$, we call *support* of T_i, and we denote it by $\mathrm{supp}(T_i)$, the subset of indices $j \subset [n]$ such that T_i depends effectively on x_j, in the sense that there is no map $h : \mathbb{R}^{n-1} \to \mathbb{R}$ such that

$$T_i(x) = h(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \ , \quad \forall x \in \mathbb{R}^n \ .$$

The following lemma provides an alternative characterization of the support.

Lemma 6.5 ([GG04, Prop. 2]). *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a convex Shapley operator. For all $i \in [n]$, index j is in $\mathrm{supp}(T_i)$ if, and only if, $\lim_{\alpha \to +\infty} T_i(\alpha e_{\{j\}}) = +\infty$.*

We now define a one-player game, denoted by Γ'_T, as follows. The state space is $[n]$ and when in state i the player, called MAX, can choose a nonempty subset of state $J \subset \mathrm{supp}(T_i)$ such that

$$\lim_{\alpha \to -\infty} T_i(\alpha e_{J^c}) > -\infty \ .$$

Once this subset is selected, the next state is chosen in J with uniform probability. For the same reason as with Γ_T, we do not need to define a payoff function.

The game Γ'_T is well defined since in any state $i \in [n]$, MAX has the possibility to choose action $J = \mathrm{supp}(T_i)$. Indeed, T_i does not depend on any x_j with $j \notin J$, which yields that $T_i(\alpha e_{J^c})$ is independent of α. Moreover, $\mathrm{supp}(T_i)$ is nonempty, otherwise T_i would be constant.

By construction, any action of MAX in Γ'_T is also an action of MAX in Γ_T. The following result gives a converse statement.

Lemma 6.6. *Let T be a convex Shapley operator on $\mathbb{R}^n$. For every state in $[n]$, if player MAX can choose a subset of states J in Γ_T, then there exists an action $K \subset J$ for MAX in Γ'_T.*

Proof. Let $J \subset [n]$ be a possible action of MAX in state i in the game Γ_T, that is, such that $\alpha \mapsto T_i(\alpha e_{J^c})$ is lower bounded. If there is some $j \in J \setminus \mathrm{supp}(T_i)$, then T_i does not depend on x_j and so $\alpha \mapsto T_i(\alpha e_{J^c \cup \{j\}})$ is still lower bounded. Thus, proceeding by induction, we find a subset $K \subset J \cap \mathrm{supp}(T_i)$ such that the sequence $\lim_{\alpha \to -\infty} T_i(\alpha e_{K^c})$ is finite. □

We define the same notion of dominion in Γ'_T as in Γ_T, which amounts to say that a subset D is a dominion if in any state $i \in D$, MAX can choose a subset $J \subset D$. In particular, Lemma 6.6 yields that a subset of states is a dominion of MAX in the two-player game Γ_T if, and only if, it is a dominion in the one-player game Γ'_T. We shall also need the following notion. We say that a nonempty subset $S \subset [n]$ is a *sink* if MAX cannot make the state to leave S with positive probability once it has reached it, that is, for any state $i \in S$, the action space of MAX contains only subsets $J \subset S$. Equivalently, S is a sink if, and only if, $\mathrm{supp}(T_i) \subset S$ for all $i \in S$. Observe that the set of sinks is not empty, since it contains the set of all states. Moreover, a sink is by definition also a dominion.

We can now adapt Theorem 6.2 to the convex case.

Theorem 6.7. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a convex Shapley operator. All slice spaces $\mathcal{S}_\alpha^\beta(T)$ are bounded in Hilbert's seminorm if, and only if, MAX has no dominion which has an empty intersection with a sink in Γ'_T.*

Proof. Suppose first that in Γ'_T there exist a sink S and a dominion D with empty intersection. We readily have that D is a dominion of MAX in Γ_T. Moreover, for every $i \in S$, we have $\mathrm{supp}(T_i) \subset S$. Hence, $T_i(\alpha e_{S^c})$ is independent of α and so, S can be chosen by MIN in the game Γ_T. This shows that S is a dominion of MIN in the latter two-player game. Thus, the dominion condition holds in Γ_T which implies that there exists a slice space unbounded in Hilbert's seminorm, according to Theorem 6.2.

Suppose now there exist an unbounded slice space, that is, according to the aforementioned theorem, players MIN and MAX have each a dominion, I and J respectively, such that $I \cap J = \emptyset$. We already know that J is also a dominion of MAX in Γ'_T. We next show that I is a sink in Γ'_T, which will conclude the proof. Let $i \in I$ and $j \notin I$. By monotonicity of T we have, for all positive scalars α, $T_i(\alpha e_{\{j\}}) \leqslant T_i(\alpha e_{I^c})$. Since the right-hand side of the latter inequality is bounded by a constant independent of α, we deduce from Lemma 6.5 that $j \notin \operatorname{supp}(T_i)$. This yields that $\operatorname{supp}(T_i) \subset I$. Since this is true for any $i \in I$, we deduce that I is a sink in Γ'_T. □

6.2 Hypergraph characterization and complexity aspects

In this section, we give a graph-theoretic construction that allows us to check the dominion condition.

6.2.1 Reachability conditions and dominions

Given any Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$, the actions in the stochastic game Γ_T are determined by the "behavior at infinity" of T along some directions. To express that behavior, we introduce a pair of hypergraphs, denoted by $(\mathcal{H}^+, \mathcal{H}^-)$, defined as follows:

- the set of nodes for both $\mathcal{H}^+$ and $\mathcal{H}^-$ is $[n]$;
- the hyperarcs of $\mathcal{H}^+$ are the pairs $(J, \{i\})$ such that $i \notin J$ and

$$\lim_{\alpha \to +\infty} T_i(\alpha e_J) = +\infty \ ;$$

- the hyperarcs of $\mathcal{H}^-$ are the pairs $(J, \{i\})$ such that $i \notin J$ and

$$\lim_{\alpha \to -\infty} T_i(\alpha e_J) = -\infty \ .$$

From the definition of the hypergraphs $\mathcal{H}^\pm$, we readily get the following characterization of invariant subsets in these hypergraphs.

Lemma 6.8. *A subset of nodes $J \subset [n]$ is invariant in $\mathcal{H}^+$ (resp., $\mathcal{H}^-$) if, and only if,*

$$\lim_{\alpha \to +\infty} T_i(\alpha e_J) < +\infty \ , \quad \forall i \in [n] \setminus J$$
$$\Big(\textit{resp.,} \ \lim_{\alpha \to -\infty} T_i(\alpha e_J) > -\infty \ , \quad \forall i \in [n] \setminus J \ \Big) \ .$$

□

The above lemma yields that a nonempty subset I is a dominion of MIN in Γ_T if, and only if, its complement is invariant in $\mathcal{H}^+$, i.e., $\operatorname{reach}(I^c, \mathcal{H}^+) = I^c$. Likewise, a nonempty subset J is a dominion of MAX in Γ_T if, and only if, its complement is invariant in $\mathcal{H}^-$, i.e., $\operatorname{reach}(J^c, \mathcal{H}^-) = J^c$. Thus, Theorem 6.2 can be reformulated in terms of reachability

Theorem 6.9. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a Shapley operator operator. All slice spaces $\mathcal{S}_\alpha^\beta(T)$ are bounded in Hilbert's seminorm if, and only if, there are no nonempty disjoint subsets of $[n]$, (I, J), such that $\operatorname{reach}(I^c, \mathcal{H}^+) = I^c$ and $\operatorname{reach}(J^c, \mathcal{H}^-) = J^c$.* □

Remark 6.10. If T satisfies Assumption 4.A, then T and its recession operator $\widehat{T}$ have the same asymptotic behaviors. Therefore, the hypergraph $\mathcal{G}^+$ (resp., $\mathcal{G}^-$) defined in Section 5.2 is the same as $\mathcal{H}^+$ (resp., $\mathcal{H}^-$). Thus, the condition in Theorem 6.9 recovers Point (ii) of Corollary 5.13.

6.2.2 Convex case

In this subsection, we consider a convex Shapley operator $T : \mathbb{R}^n \to \mathbb{R}^n$. In the same way as the dominion condition 6.C in the two-player game Γ_T simplifies to a condition in the one-player game Γ'_T, we next show that the hypergraph-reachability conditions of Theorem 6.9 have a simpler formulation.

We associate to T the directed graph, denoted by $\mathcal{G}$, with set of vertices $[n]$ and an edge from i to j if $\lim_{\alpha\to+\infty} T_i(\alpha e_{\{j\}}) = +\infty$, or equivalently, if $j \in \operatorname{supp}(T_i)$ (see Lemma 6.5). Recall that a *final class* of any directed graph is a nonempty set of nodes C such that every two nodes of C are connected by a directed path and every path starting from a node in C remains in it. Also recall that a subset S is a sink in the one-player game Γ'_T if, and only if, $\operatorname{supp}(T_i) \subset S$ for every $i \in S$. Thus, a final class of $\mathcal{G}$ is also a sink in Γ'_T and it is readily seen that any sink of the latter game contains a final class of $\mathcal{G}$.

The translation of Theorem 6.7 in terms of graph leads to the following.

Theorem 6.11. *Let T be a convex Shapley operator on $\mathbb{R}^n$. All slice spaces $\mathcal{S}_\alpha^\beta(T)$ are bounded in Hilbert's seminorn if, and only if, $\mathcal{G}$ has a unique final class C and* $\operatorname{reach}(C, \mathcal{H}^-) = [n]$.

Proof. Following Theorem 6.7, we show that in the one-player game Γ'_T, player MAX has a dominion disjoint from a sink if, and only if, the directed graph $\mathcal{G}$ has more than one final class, or a unique final class which does not have access to the whole set of nodes in $\mathcal{H}^-$.

First suppose that $\mathcal{G}$ has two distinct final classes. Then, these sets are both sinks in Γ'_T, and since any sink is also a dominion of MAX, it follows that the dominion-sink condition holds.

Next, assume that $\mathcal{G}$ has a unique final class C and that $\operatorname{reach}(C, \mathcal{H}^-) \neq [n]$. Let $D := [n] \setminus \operatorname{reach}(C, \mathcal{H}^-)$. The subsets C and D are nonempty and disjoint. Furthermore, C is a sink and D is by construction a dominion of MAX in Γ_T (since its complement is invariant in $\mathcal{H}^-$), hence it is also a dominion in Γ'_T.

Now, assume that S is a sink and D a dominion in Γ'_T such that $S \cap D = \emptyset$, and that $\mathcal{G}$ has a unique final class, denoted C. Then, we necessarily have $C \subset S$, since any sink contains a final class. Hence, C and D are disjoint, that is, $C \subset D^c$, which yields $\operatorname{reach}(C, \mathcal{H}^-) \subset \operatorname{reach}(D^c, \mathcal{H}^-) = D^c \neq [n]$. □

6.2.3 Complexity aspects

Given any Shapley operator T on $\mathbb{R}^n$, the basic issue under consideration is to check whether the ergodic equation (2.5) is solvable. Theorem 6.9 (or Theorem 6.11 in the convex case) provides a combinatorial condition for this property to hold. This condition can be effectively checked as soon as the limits $\lim_{\alpha\to\pm\infty} T_i(\alpha e_J)$ arising in the definition of the hyperarcs of $\mathcal{H}^\pm$ can be computed, which happens in general situations, see examples in Section 6.3. To set aside the latter problem, let us introduce the oracle Ω which answers "yes" to any instance $(J, i, \pm)$ if, and only if, $T_i(\alpha e_J) \to \pm\infty$ as $\alpha \to \pm\infty$. We call Turing machine with oracle Ω a Turing machine which can send a query to Ω and read the output. A call to Ω is counted as one computational step of the Turing machine. We refer the reader to [AB09] for a detailed presentation of oracle Turing machines.

We shall need the following lemma, which gives a bound for the time required to compute the set of reachable nodes in $\mathcal{H}^\pm$ from any subset. We use the notation $\operatorname{poly}(n)$ to indicate a polynomial function of n.

Lemma 6.12. *For any subset $J \subset [n]$, $\operatorname{reach}(J, \mathcal{H}^\pm)$ can be computed in $O(n^2)$ steps by a Turing machine with oracle Ω.*

Proof. Let $J_1 = J$. If $T_i(\alpha e_{J_1})$ remains bounded as $\alpha \to \pm\infty$ for all $i \notin J_1$, then J_1 is invariant in $\mathcal{H}^\pm$ according to Lemma 6.8. This means that $J_1 = \mathrm{reach}(J, \mathcal{H}^\pm)$. Otherwise, define J_2 as the union of J_1 and all the nodes $i \notin J_1$ for which $T_i(\alpha e_{J_1})$ tends to $\pm\infty$ as $\alpha \to \pm\infty$. Repeating the same steps, we arrive at a subset J_k for some integer $k \leqslant n$, which is invariant by $\mathcal{H}^\pm$ and contains J. Hence, we must have $\mathrm{reach}(J, \mathcal{H}^\pm) \subset J_k$ since $\mathrm{reach}(J, \mathcal{H}^\pm)$ is the smallest subset satisfying the latter properties. The other inclusion being trivial, we get $J_k = \mathrm{reach}(J, \mathcal{H}^\pm)$.

Now observe that each step ℓ of the latter recursion requires $|(J_\ell)^c|$ calls to the oracle Ω. So, the number of calls is bounded by n^2. Furthermore, the number of elementary operations in each step is linear with respect to the number of calls. Hence the result. □

It readily follows from the characterization of dominions that the problem of deciding whether a subset J is a dominion, i.e., $\mathrm{reach}(J^c, \mathcal{H}^\pm) = J^c$, can be solved in $O(|J|)$ steps by the oracle Turing machine. Furthermore, it is easily seen that the reachability conditions of Theorem 6.9 boil down to check that, for every $I \subset [n]$, either $\mathrm{reach}(I^c, \mathcal{H}^+) \neq I^c$ or $\mathrm{reach}(I, \mathcal{H}^-) = [n]$. Then, we readily get the following.

Theorem 6.13. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be any Shapley operator. The problem of deciding whether all slice spaces are bounded in Hilbert's seminorm can be solved in $O(2^n n^2)$ steps by a Turing machine with oracle Ω.* □

Recall that the exponential bound cannot be reduced to a polynomial one unless P = coNP. Indeed, the restricted version of this problem concerning Shapley operators with finite action spaces reduces to check the existence of a nontrivial fixed point of a payment-free Shapley operator, a problem which has been shown to be coNP-complete, see Subsection 5.4.2 and in particular Corollary 5.30. However, when T is convex, the condition in Theorem 6.11 requires the computation of the final classes of the directed graph $\mathcal{G}$. Since it has n nodes, it is known that finding its strongly connected components can be done in $O(n^2)$ time. This leads us to the following bound.

Theorem 6.14. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a convex Shapley operator. The problem of deciding whether all slice spaces are bounded in Hilbert's seminorm can be solved in $O(n^2)$ steps by a Turing machine with oracle Ω.* □

Let us point out that simpler sufficient conditions given in [GG04] (see Subsection 2.3.3), involving directed graphs instead of hypergraphs, can be checked using only a polynomial number of elementary operations, but they are less accurate.

6.3 Examples and applications

6.3.1 Shapley operator with unbounded payments

Here, we provide the example of an ergodic stochastic game for which the conditions given in [GG04, AGH15a] do not apply. We also illustrate the hypergraph characterization of Theorem 6.9. The example is an adaptation of the classical Blackmailer's Dilemma [Whi83], a one-player game with two states, one absorbing with a zero payoff and one characterized by a dilemma between obtaining a good stage payoff and increasing the probability to leave the state.

Let us consider the Shapley operator $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$T(x) = \begin{pmatrix} \sup_{0<p\leqslant 1} \big(\log p + p(x_2 \wedge x_3) + (1-p)x_1\big) \\ \inf_{0<p\leqslant 1} \big(-\log p + px_3 + (1-p)x_1\big) \\ x_3 \end{pmatrix} \enspace .$$

where $\wedge$ stands for $\min$. The map T is the dynamic programming operator of a zero-sum stochastic game with three states: the first player, called here Player I, partially controls state 1, the second player, called Player II, controls state 2, and state 3 is an absorbing state. In state 1, Player I chooses an action $p \in (0,1]$ and receives $\log p$ from Player II. Then with probability $1-p$, the next state remains 1, and with probability p it is chosen by Player II between state 2 and state 3. Thus, in order to maximize the stage payoff, one would select $p = 1$, but this leads to leave state 1 with probability one. A dual interpretation applies to Player II in state 2.

Let us denote by A_i and B_i the action spaces of players MIN and MAX, respectively, in any state i of the game Γ_T. In order to determine these spaces, it is convenient to notice that

$$T_1(x) = h\big((x_2 \wedge x_3) - x_1\big) + x_1 \quad \text{and} \quad T_2(x) = -h(x_1 - x_3) + x_1 \enspace ,$$

where h is the real function defined by

$$h(z) = \sup_{0<p\leqslant 1} (\log p + pz) \enspace , \quad \forall z \in \mathbb{R} \enspace .$$

Also note that h satisfies $h(z) = -1 - \log(-z)$ for $z \leqslant -1$ and $h(z) = z$ for $z \geqslant -1$. Thus, we get

$$\begin{aligned} A_1 &= \{\{1,2\},\{1,3\},\{1,2,3\}\} \enspace , & B_1 &= \{\{2,3\},\{1,2,3\}\} \enspace , \\ A_2 &= \{\{3\},\{1,3\},\{2,3\},\{1,2,3\}\} \enspace , & B_2 &= \{\{1,3\},\{1,2,3\}\} \enspace , \\ A_3 &= \{\{3\},\{1,3\},\{2,3\},\{1,2,3\}\} \enspace , & B_3 &= \{\{3\},\{1,3\},\{2,3\},\{1,2,3\}\} \enspace . \end{aligned}$$

Then, the dominions of MIN in Γ_T are $\{3\}$, $\{1,3\}$, $\{2,3\}$ and $\{1,2,3\}$, whereas the dominions of MAX are $\{3\}$ and $\{1,2,3\}$. It follows that the dominion condition is not satisfied, and so all slice spaces of T are bounded in Hilbert's seminorm. As a consequence, T is ergodic, that is, the ergodic equation (2.5) is solvable for all operators $g + T$ with $g \in \mathbb{R}^3$.

It is instructive to observe that MAX (resp., MIN) can choose a subset J in state i in the game Γ_T if, and only if, Player I (resp., Player II) can force the state at the next stage to be in J almost surely in the game with Shapley operator T.

In order to check the ergodicity of T, one may alternatively construct the hypergraphs $\mathcal{H}^\pm$ associated with T. Figure 6.1 shows a concise representation of these hypergraphs, in which only the hyperarcs with minimal tail (with respect to the inclusion partial order) have been represented. For instance, there is no arc from $\{2\}$ to $\{1\}$ in $\mathcal{H}^+$ since $T_1(\alpha e_2) = 0$ for all $\alpha \geqslant 0$. However, there is a hyperarc from $\{2,3\}$ to $\{1\}$, since $T_1(\alpha e_{\{2,3\}}) = \alpha$ for all $\alpha \geqslant 0$, which yields $\lim_{\alpha\to+\infty} T_1(\alpha e_{\{2,3\}}) = +\infty$. Then, one may check that there are no nonempty disjoint subsets of $[n]$, (I,J), such that $\mathrm{reach}(I^c, \mathcal{H}^+) = I^c$ and $\mathrm{reach}(J^c, \mathcal{H}^-) = J^c$, as in Theorem 6.9.

The above conclusions cannot be obtained from the conditions stated in [GG04] or in Section 4.2 in terms of fixed points of the recession operator. Indeed, here the latter operator writes

$$\widehat{T}(x) = \begin{pmatrix} x_1 \vee (x_2 \wedge x_3) \\ x_1 \wedge x_3 \\ x_3 \end{pmatrix} \enspace .$$

Hence, any vector $(\alpha, 0, 0)^\intercal$ with $\alpha \geqslant 0$ is a fixed point of $\widehat{T}$.

Figure 6.1: Hypergraphs $\mathcal{H}^+$ (left) and $\mathcal{H}^-$ (right) associated with T

6.3.2 Nonnegative tensors

Consider a d-order n-dimensional tensor $\mathcal{A}$ defined by n^d real entries, denoted by $a_{i_1\dots i_d}$ with indices $i_1, \dots, i_d \in [n]$. It yields a self-map f of $\mathbb{R}^n$, the ith coordinate function of which is given by

$$f_i(x) = (\mathcal{A}x^{(d-1)})_i := \sum_{i_2\dots i_d \in [n]} a_{i\,i_2\dots i_d}\, x_{i_2} \dots x_{i_d} \enspace .$$

An important problem, introduced independently by Lim [Lim05] and Qi [Qi05], concerns the existence of an eigenvalue $\lambda \in \mathbb{R}$ and an eigenvector $u \in \mathbb{R}^n$, solution of

$$f(u) = \mathcal{A}u^{(d-1)} = \lambda u^{d-1} \enspace , \tag{6.3}$$

where $u^{d-1} = (u_1^{d-1}, \dots, u_n^{d-1})$. If the tensor $\mathcal{A}$ is nonnegative, meaning that $a_{i_1\dots i_d} \geqslant 0$ for all indices, a variant of this problem is the existence of a positive eigenvalue $\lambda > 0$ and an eigenvector $u \in \operatorname{int}(\mathbb{R}_+^n)$ in the standard positive cone of $\mathbb{R}^n$. The latter problem can be seen as a particular case of a nonlinear Perron-Frobenius eigenproblem. In particular, Friedland, Gaubert and Han [FGH13] showed, as a consequence of [GG04, Th. 2], that a nonnegative tensor $\mathcal{A}$ has a positive eigenvalue $\lambda > 0$ and an eigenvector $u \in \operatorname{int}(\mathbb{R}_+^n)$ if $\mathcal{A}$ is *weakly irreducible*. This notion can be defined by means of the directed graph $\mathcal{G}(\mathcal{A})$, with set of nodes $[n]$ and an edge from i to j if there exists a set of indices $i_2, \dots, i_d$ containing j and such that $a_{i\,i_2\dots i_d} > 0$. Then, the nonnegative tensor $\mathcal{A}$ is *weakly irreducible* if $\mathcal{G}(\mathcal{A})$ is strongly connected.

Theorem 6.11 allows us to extend the latter result. For that purpose, let us introduce the hypergraph $\mathcal{H}(\mathcal{A})$ with set of nodes $[n]$ and an hyperarc from $J \subset [n]$ to $\{i\}$ if $i \notin J$ and $a_{i\,i_2\dots i_d} > 0$ implies that $J \cap \{i_2, \dots, i_d\} \neq \emptyset$.

Corollary 6.15. *Let $\mathcal{A}$ be a nonnegative tensor of order d and dimension n. If the directed graph $\mathcal{G}(\mathcal{A})$ has a unique final class C and if $\operatorname{reach}(C, \mathcal{H}(\mathcal{A})) = [n]$, then $\mathcal{A}$ has a positive eigenvalue and an eigenvector in $\operatorname{int}(\mathbb{R}_+^n)$.*

Proof. Consider $T := (d-1)^{-1} \log \circ f \circ \exp$. This is a monotone additively homogeneous self-map of $\mathbb{R}^n$. Furthermore, any eigenpair $(\mu, v) \in \mathbb{R} \times \mathbb{R}^n$ of T yields an eigenpair (λ, u) of f with the required properties, namely $\lambda = e^{\mu(d-1)} > 0$ and $u = \exp(v) \in \operatorname{int}(\mathbb{R}_+^n)$. This is a standard result that functions such as $\log \circ f_i \circ \exp$ are convex – it relies on the convexity of the "log-exp" function $x \mapsto \log(e^{x_1} + \dots + e^{x_n})$, see [RW98, Ex. 2.16, Ex. 2.52]. Hence T is convex. Now, observe that the directed graph $\mathcal{G}$ and the hypergraph $\mathcal{H}^-$ associated with the convex Shapley operator T are the same as $\mathcal{G}(\mathcal{A})$ and $\mathcal{H}(\mathcal{A})$, respectively. The conclusion follows from an application of Theorem 6.11. □

Part III

Generic uniqueness of the bias vector of zero-sum stochastic games

CHAPTER 7

Generic uniqueness of the bias of perfect-information finite stochastic games

7.1 Introduction

The description of the set of bias vectors of zero-sum stochastic games is a fundamental issue. For one-player problems, i.e., for discrete optimal control, the representation of bias vectors and their relation with optimal strategies is well understood, in either the deterministic or the stochastic case (MDPs). In the deterministic case, the analysis relies on max-plus spectral theory, which goes back to the work of Romanovsky [Rom67], Gondran and Minoux [GM77] and Cuninghame-Green [CG79]. We also refer the reader to [MS92, BCOQ92, ABG13] for more background on max-plus spectral theory. Kontorer and Yakovenko [KY92] and Kolokoltsov and Maslov [KM97] deal specially with infinite-horizon optimization and mean-payoff problems. In this framework, the set of bias vectors has the structure of a max-plus (tropical) cone, i.e., it is invariant by max-plus linear combinations, and it has a unique minimal generating family consisting of certain "extreme" generators, which can be identified by looking at the support of the maximizing measures in the linear programming formulation of the optimal control problem, or at the "recurrence points" of infinite optimal trajectories. A combinatorial interpretation of some of these results, in terms of polyhedral fans, has been recently given by Sturmfels and Tran [ST13]. The ergodic equation (2.5) and the structure of bias vectors has also been studied for an infinite-dimensional state space in the context of infinite-dimensional max-plus spectral theory, see Akian, Gaubert and Walsh [AGW09], and also in the setting of weak KAM theory, for which we refer the reader to Fathi [Fat08]. In the stochastic case, the structure of the set of bias vectors is also known, at least when the state space is finite, see Akian and Gaubert [AG03].

In the two-player case, the structure of the set of bias vectors is less well known. In particular, the uniqueness of the bias vector up to an additive constant is an important matter for algorithmic purposes. This is specially the case when applying the standard Hoffman-Karp policy iteration algorithm, since nonuniqueness typically leads to numerical instabilities or degeneracies. Hoffman and Karp [HK66] introduced this algorithm to solve mean-payoff zero-sum stochastic games with perfect information and finite state and action spaces. They showed that policy iteration does terminate if every pair of strategies of the two players yields an irreducible Markov chain. However, policy iteration may cycle if this irreducibility assumption is not satisfied, which is the case for many classes of games – in particular, it is essentially never satisfied for deterministic games. Some refinements of the Hoffman-Karp scheme have been proposed by Cochet-Terrasson and Gaubert [CTG06], Akian, Cochet-Terrasson, Detournay and Gaubert [ACTDG12], Bourque and Raghavan [BR14], allowing one to circumvent such degeneracies at the price of an increased complexity of the algorithm (handling the nonuniqueness of the bias vector). Hence, it is of interest to understand when such technicalities can be avoided.

In this chapter, we address the question of the uniqueness of the bias vector of stochastic games with perfect information and finite state and action spaces, restricting our attention to ergodic games, i.e., for which the ergodic equation (2.5) is solvable for all (state-dependent) perturbations of the transition payments. The main result, Theorem 7.8, shows that the bias vector is generically unique up to an additive constant. More precisely, it shows that the set of perturbation vectors for which the bias vector is not unique belongs to a polyhedral complex the cells of which have codimension one at least. We then deduce that the Hoffman-Karp policy iteration algorithm does converge if the payment is generic. This leads to an explicit perturbation scheme, allowing one to solve nongeneric instances by policy iteration, avoiding the classical irreducibility condition.

In the following chapter, we address the generic uniqueness problem for any zero-sum stochastic game with finite state space. We show that generic uniqueness of the bias vector holds locally, that is, for a generic perturbation in a closed set with nonempty interior, as soon as the ergodic equation (2.5) is solvable for all perturbations in a neighborhood of this set.

The theoretic results (Section 7.3) persented in this chapter have been announced in the CDC conference proceedings [AGH14b]. The application to policy iteration (Section 7.4) is new.

7.2 Further preliminaries in nonlinear Perron-Frobenius theory

In the remainder of this chapter, we consider a finite zero-sum stochastic game with perfect information $\Gamma = ([n], A, B, K_A, K_B, r, p)$. Since the action spaces A and B are finite, its Shapley operator is a piecewise affine map, the ith coordinate of which writes

$$T_i(x) = \min_{a \in A_i} \max_{b \in B_{i,a}} \Big(r(i,a,b) + \sum_{j \in [n]} x_j \, p(j \mid i,a,b) \Big) \ , \quad \forall x \in \mathbb{R}^n \ . \tag{7.1}$$

Recall in particular that T has an invariant half-line (Theorem 2.5), that is, there exist two vectors $u, \nu \in \mathbb{R}^n$ such that $T(u + \alpha\nu) = u + (\alpha+1)\nu$ for all $\alpha \geqslant 0$. Furthermore, the vector ν is unique and equal to the mean-payoff vector of Γ: $\chi(T) = \lim_{k\to\infty} T^k(0)/k$.

7.2.1 Characterization of the upper mean payoff

Let us denote by $\mathcal{S}_s$ the finite set of deterministic stationary strategies of player MIN, also called (deterministic) policies. It is the set of maps $\sigma : [n] \to A$ such that $\sigma(i) \in A_i$ for every state $i \in [n]$. Likewise, we denote by $\mathcal{T}_s^\sigma$ the set of deterministic stationary strategies of player MAX when the policy σ has been chosen by MIN, i.e., the set of maps $\tau : [n] \to B$ such that $\tau(i) \in B_{i,\sigma(i)}$ for every state $i \in [n]$.

We introduce the reduced Shapley operator $T^\sigma : \mathbb{R}^n \to \mathbb{R}^n$ associated with the policy $\sigma \in \mathcal{S}_s$ of player MIN. Its ith coordinate map is given by

$$T_i^\sigma(x) = \max_{b \in B_{i,\sigma(i)}} \Big(r(i,\sigma(i),b) + \sum_{j \in [n]} x_j \, p(j \mid i,\sigma(i),b) \Big) \ , \quad \forall x \in \mathbb{R}^n \ .$$

We endow $\mathbb{R}^n$ with the usual partial order. Then, since the action spaces are finite, we readily have, for all $x \in \mathbb{R}^n$,

$$T(x) = \min_{\sigma \in \mathcal{S}_s} T^\sigma(x) \ . \tag{7.2}$$

Likewise, we have, for all $x \in \mathbb{R}^n$,

$$T^\sigma(x) = \max_{\tau \in \mathcal{T}_s^\sigma} \left(r^{\sigma\tau} + P^{\sigma\tau} x \right) \ , \tag{7.3}$$

where $r^{\sigma\tau}$ is the vector in $\mathbb{R}^n$ whose ith entry is defined by

$$r_i^{\sigma\tau} = r(i, \sigma(i), \tau(i))$$

and $P^{\sigma\tau}$ is the $n \times n$ stochastic matrix whose (i,j) entry is given by

$$P_{ij}^{\sigma\tau} = p(j \mid i, \sigma(i), \tau(i)) \ .$$

Observe that T^σ is monotone, additively homogeneous and convex, in the sense that each of its coordinates is a convex function.

Recall that if P is a $n \times n$ stochastic matrix, the directed graph associated with P is composed of the nodes $1, \dots, n$ and of the arcs (i,j) such that $P_{ij} > 0$. A class of the matrix P is a maximal set of nodes such that every two nodes in this set are connected by a directed path. A class is said to be final if every path starting from a node of this class remains in it. Let us denote by $\mathcal{M}(P)$ the set of invariant probability measures of P, i.e., the set of stochastic vector $m \in \mathbb{R}^n$ such that $m^\intercal P = m^\intercal$. Given a final class C of P, there is a unique invariant probability measure $m \in \mathcal{M}(P)$ the support of which is C, i.e., $\{i \in [n] \mid m_i > 0\} = C$. Moreover, the set $\mathcal{M}(P)$ is the convex hull of such measures, and since the number of final class of P is finite, $\mathcal{M}(P)$ is a convex polytope.

Let us denote by $\overline{\chi}(T)$ the *upper mean payoff* of T, i.e., the greatest entry of the mean-payoff vector $\chi(T)$. We next give a characterization of $\overline{\chi}(T)$. Obviously, if the ergodic equation (2.5) is solvable, then $\chi(T)$ is a constant vector, the coordinates of which are equal to the eigenvalue $\lambda(T)$, so that the following result also provides a characterization of the eigenvalue. In the sequel, we denote by $\langle x, y \rangle$ the standard scalar product in $\mathbb{R}^n$ of two vectors x, y.

Lemma 7.1. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator (7.1) of a finite stochastic game with perfect information Γ. Then the upper mean payoff of T is given by*

$$\overline{\chi}(T) = \min_{\sigma \in \mathcal{S}_s} \max \{ \langle m, r^{\sigma\tau} \rangle \mid \tau \in \mathcal{T}_s^\sigma, \ m \in \mathcal{M}(P^{\sigma\tau}) \} \ . \tag{7.4}$$

Proof. First, observe that for all policies $\sigma \in \mathcal{S}_s$ we have $T \leqslant T^\sigma$, which yields, by monotonicity of the operators, $\chi(T) \leqslant \chi(T^\sigma)$ and in particular $\overline{\chi}(T) \leqslant \overline{\chi}(T^\sigma)$.

Considering an invariant half-line of T, we know that there exist a vector $u \in \mathbb{R}^n$ such that $T(u) = u + \chi(T)$. Let $\sigma \in \mathcal{S}_s$ be a policy of player MIN such that $T(u) = T^\sigma(u)$. Then, we have $T^\sigma(u) \leqslant u + \overline{\chi}(T) e$. Furthermore, we know by a Collatz-Wielandt formula (see [GG04, Prop. 1]) that, for any monotone and additively homogeneous map $F : \mathbb{R}^n \to \mathbb{R}^n$, we have

$$\overline{\chi}(F) = \inf \{ \mu \in \mathbb{R} \mid \exists x \in \mathbb{R}^n, \ F(x) \leqslant \mu e + x \} \ .$$

So, we deduce that $\overline{\chi}(T^\sigma) \leqslant \overline{\chi}(T)$, and finally that

$$\overline{\chi}(T) = \min_{\sigma \in \mathcal{S}_s} \overline{\chi}(T^\sigma) \ .$$

Now we fix a policy σ of MIN, and we let $\chi := \chi(T^\sigma)$ and

$$\mu^\sigma := \max\{\langle m, r^{\sigma\tau}\rangle \mid \tau \in \mathcal{T}_s^\sigma,\ m \in \mathcal{M}(P^{\sigma\tau})\}\ .$$

Since T^σ has an invariant half-line with direction χ, there exists a vector $v \in \mathbb{R}^n$ such that $T^\sigma(v + \alpha\chi) = v + (\alpha+1)\chi$ for all $\alpha \geqslant 0$. In particular, for every policy $\tau \in \mathcal{T}_s^\sigma$ we have

$$r^{\sigma\tau} + P^{\sigma\tau} v \leqslant T^\sigma(v) = v + \chi \leqslant v + \overline{\chi}(T^\sigma) e\ .$$

Then, we easily deduce that $\mu^\sigma \leqslant \overline{\chi}(T^\sigma)$.

Furthermore, since the germs of affine functions from $\mathbb{R}$ to $\mathbb{R}$ at infinity are totally ordered, there exists a policy $\tau \in \mathcal{T}_s^\sigma$ such that

$$T^\sigma(v + \alpha\chi) = r^{\sigma\tau} + P^{\sigma\tau}(v + \alpha\chi)$$

for all α large enough. In particular, since the equality

$$v + (\alpha+1)\chi = r^{\sigma\tau} + P^{\sigma\tau}(v + \alpha\chi) \tag{7.5}$$

holds for all α large enough, we get that $P^{\sigma\tau}\chi = \chi$. Thus, χ is an *harmonic vector* for the stochastic matrix $P^{\sigma\tau}$. As such, it is constant on any final class of $P^{\sigma\tau}$ and its maximum is attained on one of these final class (see [AG03, Lem. 2.9]). Let $m \in \mathcal{M}(P^{\sigma\tau})$ be the invariant probability measure associated with a final class C of $P^{\sigma\tau}$ such that $\chi_i = \overline{\chi}(T^\sigma)$ for all $i \in C$. Then, we deduce from (7.5) that $\langle m, r^{\sigma\tau}\rangle = \overline{\chi}(T^\sigma)$, which yields $\mu^\alpha \geqslant \overline{\chi}(T^\sigma)$ and finally $\mu^\alpha = \overline{\chi}(T^\sigma)$. □

Let us mention that in (7.4), the set of invariant probability measures of $P^{\sigma\tau}$, $\mathcal{M}(P^{\sigma\tau})$, may be replaced by the set of its extreme points, denoted by $\mathcal{M}^*(P^{\sigma\tau})$, since it is a convex polytope. Recall that $\mathcal{M}^*(P^{\sigma\tau})$ is the set of invariant probability measures the support of which are the final classes of $P^{\sigma\tau}$.

7.2.2 Structure of the eigenspace

An ingredient of our approach is a result of [AG03] which describes the eigenspace, i.e., the set of bias vectors, of *one-player* Shapley operators T. We next recall this result. Let us assume that T arises from a game in which only player, MAX, has nontrivial actions, whereas player MIN has only one possible policy σ. In this case, the representation of the eigenvalue $\lambda(T) = \overline{\chi}(T)$ in Lemma 7.1 simplifies as the dependency in σ can be dropped, and we arrive, with a trivial simplification of the notation, to

$$\lambda(T) = \max\{\langle m, r^\tau\rangle \mid \tau \in \mathcal{T}_s,\ m \in \mathcal{M}(P^\tau)\}\ . \tag{7.6}$$

It is shown in [AG03] that the dimension of the eigenspace of T is controlled by the number of *critical classes*. The latter can be defined through the notion of maximizing measures, which rely on *behavioral* policies. Recall that for every state i, such a policy τ assigns to every action $b \in B_i$ a probability $\tau(b \mid i)$ that this action is selected. This leads to the stochastic matrix P^τ with entries $P^\tau_{ij} = \sum_{b\in B_i} p(j \mid i, b)\,\tau(b \mid i)$, and to the payment vector r^τ with entries $r^\tau_i = \sum_{b\in B_i} r(i,b)\,\tau(b \mid i)$. Then, the maximum in (7.6) is unchanged if it is taken over the set of behavioral policies τ and of invariant measures m of the corresponding stochastic matrix P^τ (see [AG03, Prop. 7.2]). Then a measure m is *maximizing* if $\lambda(T) = \langle m, r^\tau\rangle$ for some behavioral policy τ and if $m \in \mathcal{M}^*(P^\tau)$, that is, the support of m is a final class of P^τ. A subset $I \subset [n]$ is a *critical class* if there exists a maximizing measure m whose support is I and if I is a maximal element with respect to inclusion among all the subsets of $[n]$ which arise in this way. Note that critical classes are disjoint.

The following lemma gives a sufficient condition for the critical class to be unique. Note that the condition does not require behavioral stationary strategies.

Lemma 7.2. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a convex Shapley operator. Suppose that the ergodic equation (2.5) is solvable. If there is a unique probability measure which attains the maximum in (7.6), then T has a unique critical class.*

Proof. Let C be a critical class of T. There exists a behavioral policy τ such that C is a final class of P^τ and such that the invariant measure m with support C satisfies $\lambda := \lambda(T) = \langle m, r^\tau \rangle$.

Let u be an eigenvector of T. For every $i \in [n]$, we have $\lambda + u_i - r_i^\tau - P_i^\tau u \geqslant 0$. Furthermore, since $m^\intercal P^\tau = m^\intercal$, we also have

$$\sum_{i \in C} m_i \left(\lambda + u_i - r_i^\tau - P_i^\tau u\right) = \lambda + \langle m, u \rangle - \langle m, r^\tau \rangle - \langle m, (P^\tau u) \rangle = 0 \ .$$

This yields that $\lambda + u_i - r_i^\tau - P_i^\tau u = 0$ for every $i \in C$. With a similar argument, we show that $\lambda + u_i - r_i^b - P_i^b u = 0$ for every $i \in C$ and every b such that $\tau(b \mid i) > 0$.

Let τ' be a deterministic policy such that for all indices i, $\tau'(i) = b_i \in \operatorname{supp}(\tau(\cdot \mid i))$, i.e., $\tau(b_i \mid i) > 0$. Since P_i^τ is a convex combination of vectors one of which is $P_i^{\tau'}$, we deduce that C contains a final class C' of $P^{\tau'}$. Let m' be the invariant measure of $P^{\tau'}$ whose support is C'. Then, we have $\lambda = \langle m', r^{\tau'} \rangle$, which implies that m' is the unique measure attaining the maximum in (7.6). It follows that that C is necessarily unique since it must contain the support of m'. □

We now describe the eigenspace of T.

Theorem 7.3 ([AG03, Th. 1.1])**.** *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a convex Shapley operator. Suppose that the ergodic equation (2.5) is solvable, and let $C \subset [n]$ be the union of all critical classes. We denote by π_C the restriction map $\mathbb{R}^n \to \mathbb{R}^C$, $x \mapsto (x_i)_{i \in C}$. Then, the set of eigenvectors of T, denoted by $\mathcal{E}(T)$, satisfies the following properties:*

(i) every element x of $\mathcal{E}(T)$ is uniquely determined by its restriction $\pi_C(x)$;

(ii) the set $\pi_C(\mathcal{E}(T))$ is convex and its dimension is at most equal to the number of critical classes of T; moreover, the latter bound is attained when T is piecewise affine.

In particular, combined with Lemma 7.2, the above result yields the following.

Corollary 7.4. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a convex Shapley operator. Suppose that the ergodic equation (2.5) is solvable. If there is a unique probability measure which attains the maximum in (7.6), then T has a unique eigenvector up to an additive constant.*

We refer the reader to [AG03] for more background on critical classes, which admit several characterizations and can be computed in polynomial time when the game is finite. We only provide here a simple illustration in order to understand the latter theorem.

Example 7.5. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be such that

$$T(x) = \begin{pmatrix} x_1 \ \vee \ \frac{1}{2}(x_1 + x_2) \\ -3 + x_2 \ \vee \ \frac{1}{2}(x_1 + x_2) \end{pmatrix} \ .$$

We have $T(0) = 0$, which shows in particular that the upper mean payoff is 0. If MAX chooses, when in state 1, the action corresponding to the first term in the expression of T_1, and when in state 2, the action corresponding to the second term in the expression of T_2, then we arrive at the transition matrix

$$P = \begin{pmatrix} 1 & 0 \\ 1/2 & 1/2 \end{pmatrix}$$

which has the invariant measure $m = (1, 0)$. This measure attains the maximum in (7.6). However, its support $I = \{1\}$ is not a critical class, for it is not maximal with respect to inclusion.

Indeed, if MAX chooses instead, when in state 1, the action corresponding to the second term in the expression of T_1, we arrive at the transition matrix

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

which has the invariant measure $m = (1/2, 1/2)$. This measure attains the maximum in (7.6) and its support $I = \{1, 2\}$ is maximal with respect to inclusion. Hence, $I = \{1, 2\}$ is the unique critical class. It follows from Theorem 7.3 that 0 is the unique eigenvector of T, up to an additive constant.

Another ingredient is a variant of a result of Bruck [Bru73] concerning the topology of fixed-point sets of nonexpansive maps. We now consider a *two-player* Shapley operator T such that the ergodic equation (2.5) is solvable, and denote by

$$\mathcal{E}(T) := \{u \in \mathbb{R}^n \mid T(u) = \lambda e + u\}$$

the set of eigenvectors of T.

Theorem 7.6 (Compare with [Bru73, Th. 2])**.** *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a Shapley operator. Assume that the ergodic equation* (2.5) *is solvable. Then, the set of eigenvectors $\mathcal{E}(T)$ is a retract of $\mathbb{R}^n$ by a sup-norm nonexpansive map, meaning that $\mathcal{E}(T) = \pi(\mathbb{R}^n)$ where π is a sup-norm nonexpansive self-map of $\mathbb{R}^n$ such that $\pi = \pi^2$. In particular, $\mathcal{E}(T)$ is arcwise connected.*

Proof. The result of Bruck [Bru73, Th. 2] shows that, under some compactness conditions, the fixed-point set of a nonexpansive self-map of a Banach space is a retract of the whole space by a nonexpansive map.

Assume now that T is a Shapley operator and admits an eigenvector for the eigenvalue λ. Then, the eigenspace $\mathcal{E}(T)$ coincides with the fixed-point set of the map $x \mapsto -\lambda e + T(x)$. The latter map is sup-norm nonexpansive and satisfies the conditions of [Bru73, Th. 2], and so $\mathcal{E}(T)$ is a nonexpansive retract of $\mathbb{R}^n$. □

Remark 7.7. The retraction π in Theorem 7.6 can be chosen to be monotone and additively homogeneous. This can actually be shown by elementary means, following a construction in the proof of [GG04, Lem. 3]. Indeed, we may assume without loss of generality that $\lambda = 0$, and consider $q(x) := \lim_{k\to\infty} \inf_{\ell \geqslant k} T^{\ell}(x)$, which is finite because every orbit of a nonexpansive map that admits a fixed point must be bounded. Since T is monotone and continuous, we get $T(q(x)) \leqslant q(x)$, and so $\pi(x) := \lim_{k\to\infty} T^k(q(x))$, which is the limit of a nonincreasing and bounded sequence, exists and is finite. The map π is easily shown to be monotone and additively homogeneous and to satisfy $\pi = \pi^2$.

7.3 Generic uniqueness of the eigenvector

Let us first recall some definitions. A *polyhedron* in $\mathbb{R}^n$ is an intersection of finitely many half-spaces, a *face* of a polyhedron is an intersection of this polyhedron with a supporting half-space, and a *polyhedral complex* is a finite set $\mathcal{K}$ of polyhedra satisfying the following two properties:
(i) $P \in \mathcal{K}$ and F is a face of P implies that $F \in \mathcal{K}$;
(ii) for all $P, Q \in \mathcal{K}$, $P \cap Q$ is a face of P and Q.
A polyhedron in $\mathcal{K}$ is called a *cell* of the polyhedral complex. We refer to the textbook [DLRS10] for background on polyhedral complexes. Also recall that a real map on $\mathbb{R}^n$ is piecewise affine if, and only if, $\mathbb{R}^n$ can be covered by a finite union of polyhedra (with nonempty interior) on which its restriction is affine. If the latter map is continuous, then the set of such polyhedra can

be refined in a polyhedral complex. Finally, recall that continuous piecewise affine functions are exactly the maps whose coordinate functions can be written as a minimax over finite sets of affine functions.

We now state the main result of this chapter.

Theorem 7.8. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of a finite stochastic game with perfect information. Assume that T is ergodic. Then, the space $\mathbb{R}^n$ can be covered by a polyhedral complex such that, for any additive perturbation vector $g \in \mathbb{R}^n$ in the interior of a full-dimensional cell, $g + T$ has a unique eigenvector, up to an additive constant.*

In particular, the set of perturbation vectors g for which $g + T$ has more than one eigenvector, up to an additive constant, is included in the finite union of subspaces of codimension at least 1.

Remark 7.9. This perturbation theorem bears some conceptual similarity with results of weak KAM theory – we refer to the monograph by Fathi [Fat08] for more information. The latter theory deals with a class of one-player deterministic games with continuous time and space. In this setting, the eigenvector u and the eigenvalue λ are solution of an ergodic Hamilton-Jacobi PDE $H(x, D_x u) = \lambda$ where the Hamiltonian $(x, p) \mapsto H(x, p)$ is convex in the adjoint variable p. One may consider the perturbation of a Hamiltonian by a potential, which amounts to replacing $H(x, p)$ by $H(x, p) + V(x)$ for some function V. This is similar to the replacement of the Shapley operator T by the perturbed Shapley operator $g + T$ in Theorem 7.8. As observed by Figalli and Rifford in [FR13, Th. 4.2], it follows from weak KAM theory results that under some assumptions, the solution u of $V(x) + H(u, D_x u) = \lambda$ is unique up to an additive constant for a generic function V. Theorem 7.8 shows that an analogous property is valid for finite two-player zero-sum stochastic games. We note however that Theorem 7.8 does not extend easily to the case of PDE, since zero-sum games correspond to Hamilton-Jacobi PDE with a *nonconvex* Hamiltonian (to which current weak KAM methods do not apply).

Proof. Let T be the Shapley operator of a finite stochastic game with perfect information Γ which is assumed to be ergodic. Let $\sigma \in \mathcal{S}_s$ be a policy of player MIN. We define the real map $\lambda^\sigma(\cdot)$ on $\mathbb{R}^n$ by

$$\lambda^\sigma(g) := \max\{\langle m, (g + r^{\sigma\tau})\rangle \mid \tau \in \mathcal{T}_s^\sigma,\ m \in \mathcal{M}^*(P^{\sigma\tau})\} \ , \tag{7.7}$$

where $\mathcal{M}^*(P)$ denotes the set of extreme points of the convex polytope $\mathcal{M}(P)$, that is, the set of invariant probability measures the support of which are final classes of the stochastic matrix P. The fact that $\mathcal{M}^*(P^{\sigma\tau})$ is a set of probability measures yields that λ^σ is monotone and additively homogeneous, hence sup-norm nonexpansive (and continuous). Furthermore, since the set of policies $\mathcal{T}_s^\sigma$ of player MAX is finite, as well as all the sets $\mathcal{M}^*(P^{\sigma\tau})$, then the map λ^σ is piecewise affine. We now define the polyhedral complex $\mathcal{C}^\sigma$ covering $\mathbb{R}^n$, the full-dimensional cells of which are precisely the maximal polyhedra on which the piecewise affine map λ^σ coincides with a unique affine map. Therefore, if Q is a cell of $\mathcal{C}^\sigma$ with full dimension, then there is a unique vector $m \in \mathbb{R}^n$ and a unique scalar $d \in \mathbb{R}$ such that $\lambda^\sigma(g) = \langle m, g\rangle + d$ in the interior of Q. Observe that m must be a stochastic vector since the map $g \mapsto \lambda^\sigma(g)$ is monotone and additively homogeneous.

We claim that if a vector g is in the interior of Q, then the set of eigenvectors of the reduced one-player Shapley operator $F := g + T^\sigma$ is either empty or reduced to a line. To see this, it suffices to observe that the measure m attaining the maximum in (7.7) is unique for all g in the interior of Q and independent of the choice of g in this interior, because $g \mapsto \lambda^\sigma(g)$ is affine on the interior of Q and m must coincide with the linear part of this affine map. We deduce from Corollary 7.4 that $\mathcal{E}(F)$ is either empty or reduced to a line of direction e.

Consider now the polyhedral complex $\mathcal{C}$ obtained as the refinement of all the complexes $\mathcal{C}^\sigma$. This complex covers $\mathbb{R}^n$. Let g be a perturbation vector in the interior of a full-dimensional cell of $\mathcal{C}$. Since the game Γ is ergodic, $\mathcal{E}(g+T)$ is not empty. Let u be an eigenvector of $g+T$. According to (7.2), there is a policy $\sigma \in \mathcal{S}_s$ of player MIN such that $g+T(u) = g+T^\sigma(u)$. Hence u is also an eigenvector of $g+T^\sigma$. So, there is a finite family Σ^* of $\mathcal{S}_s$ such that $\mathcal{E}(g+T) = \bigcup_{\sigma\in\Sigma^*}\mathcal{E}(g+T^\sigma)$. Moreover, we have proved that for any policy $\sigma\in\Sigma^*$, the eigenspace $\mathcal{E}(g+T^\sigma)$ is reduced to a line. Thus, $\mathcal{E}(g+T)$ is composed of a finite union of lines which all have the same direction, namely e. Consider the hyperplane orthogonal to the unit vector, $H := \{x \in \mathbb{R}^n \mid \langle x, e\rangle = 0\}$, and let π denote the orthogonal projection on H. Then, $\pi(\mathcal{E}(g+T)) = \mathcal{E}(g+T)\cap H$ is finite. However, by Theorem 7.6, $\mathcal{E}(g+T)$ is connected. Then, the set $\pi(\mathcal{E}(g+T))$ is also connected, and since it is finite, it must be reduced to a point. It follows that $g+T$ has a unique eigenvector, up to an additive constant. □

Example 7.10. Consider the following Shapley operator defined on $\mathbb{R}^3$ – we use $\wedge$ and $\vee$ instead of $\min$ and $\max$, respectively, and we recall that the addition has precedence over the latter operators:

$$T(x) = \begin{pmatrix} \frac{1}{2}(x_1+x_3)\ \wedge\ 1+\frac{1}{2}(x_1+x_2) \\ 2+\frac{1}{2}(x_1+x_3)\ \wedge\ \left(1+\frac{1}{2}(x_1+x_2)\ \vee\ -2+x_3\right) \\ 3+\frac{1}{2}(x_1+x_3)\ \vee\ 1+x_3 \end{pmatrix} .$$

It can be proved, using Theorem 4.6, that T is ergodic. Figure 7.1 shows the intersection of the hyperplane $\{g\in\mathbb{R}^3 \mid g_3 = 0\}$ with the polyhedral complex introduced in Theorem 7.8. Here, for each vector g in the interior of a full-dimensional polyhedron, $g+T$ has a unique eigenvector up to an additive constant.

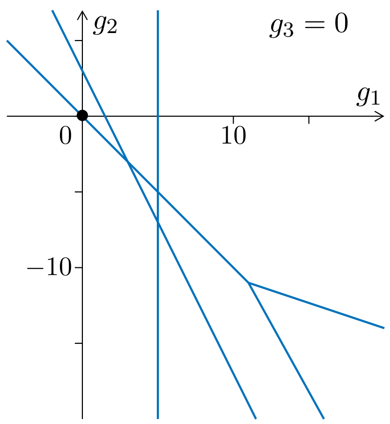

Figure 7.1: Polyhedral complex $\mathcal{C}$

Let us detail what happens in the neighborhood of $g=0$, point in which $g+T$ fails to have a unique eigenvector. Note that in the neighborhood of $g=0$, the eigenvalue of $g+T$ remains equal to 1.

- If $g_1+g_2 = 0$, the eigenvectors of $g+T$ are defined by

$$x_1 = x_2 + 2g_1 \ , \quad -3+g_2 \leqslant x_2 - x_3 \leqslant -2-g_1 \ .$$

- If $g_1+g_2 > 0$, the unique eigenvector, up to an additive constant, is

$$(-2+2g_1, -2+2g_1+2g_2, 0) \ .$$

- If $g_1+g_2 < 0$, the unique eigenvector, up to an additive constant, is

$$(-3+2g_1+g_2, -3+g_2, 0) \ .$$

7.4 Application to policy iteration

In this section, we apply the previous result to show that policy iteration combined with a perturbation scheme can solve degenerate instances of stochastic games.

7.4.1 Hoffman-Karp policy iteration

Let us recall the notation of Section 7.3. We denote by Γ a finite stochastic game with perfect information. The (finite) set of policies of player MIN is denoted by $\mathcal{S}_s$, an element of $\mathcal{S}_s$ being a map $\sigma : [n] \to A$ such that $\sigma(i) \in A_i$ for every state $i \in [n]$. The (finite) set of policies of player MAX when policy $\sigma \in \mathcal{S}_s$ of MIN is fixed is denoted by $\mathcal{T}_s^\sigma$, an element of $\mathcal{T}_s^\sigma$ being a map $\tau : [n] \to B$ such that $\tau(i) \in B_{i,\sigma(i)}$ for every state $i \in [n]$. Finally, recall that for $\sigma \in \mathcal{S}_s$ and $\tau \in \mathcal{T}_s^\sigma$, $P^{\sigma\tau}$ denotes the $n \times n$ stochastic matrix whose (i, j) entry is given by $P_{ij}^{\sigma\tau} = p(j \mid i, \sigma(i), \tau(i))$.

When $T : \mathbb{R}^n \to \mathbb{R}^n$ is the Shapley operator (7.1) of a finite stochastic game with perfect information, Hoffman and Karp [HK66] introduced a policy iteration algorithm, which takes the description of the game as input and returns the eigenvalue λ and an eigenvector u of T, i.e., a solution $(\lambda, u) \in \mathbb{R} \times \mathbb{R}^n$ of the ergodic equation $T(u) = \lambda e + u$. Also, optimal stationary strategies for both players can be derived from the output of the algorithm. It is convenient here to state an abstract, slightly more general, version of the Hoffman-Karp algorithm, described in terms of the operators T and T^σ (Algorithm 2).

Algorithm 2: Policy iteration – compare with [HK66]

input : perfect-information finite stochastic game $([n], A, B, K_A, K_B, r, p)$ with Shapley operator T.
output: eigenvalue λ and eigenvector u of T.

1 **initialization**: select an arbitrary policy $\sigma_0 \in \mathcal{S}_s$
2 **repeat**
3 compute an eigenpair (λ^k, v^k) of T^{σ_k}
4 improve the policy σ_k in a conservative way: select a policy $\sigma_{k+1} \in \mathcal{S}_s$ such that $T(v^k) = T^{\sigma_{k+1}}(v^k)$ and satisfying, for every state $i \in [n]$, $\sigma_{k+1}(i) = \sigma_k(i)$ if $T_i(v^k) = T_i^{\sigma_k}(v^k)$
5 **until** $\sigma_{k+1} = \sigma_k$
6 **return** λ^k and v^k

We assume Algorithm 2 is interpreted in exact arithmetics (the vectors v^k have rational coordinates and the λ^k are rational numbers). To implement Step 3, we may call any oracle able to compute the eigenvalue and an eigenvector of a one-player stochastic game. In the original approach of Hoffman and Karp, the oracle consists in applying the same policy iteration algorithm for the one-player game with fixed policy σ_k. The proof of Hoffman and Karp shows that Algorithm 2 is valid under a restrictive assumption.

Theorem 7.11 (Cor. of [HK66])**.** *Algorithm 2 terminates and is correct if for all choices of policies σ and τ of the two players, the corresponding transition matrix $P^{\sigma\tau}$ is irreducible.*

Indeed, it is easy to see that the sequence $(\lambda^k)_k$ of Algorithm 2 is nonincreasing, that is, $\lambda^{k+1} \leqslant \lambda^k$ for all iterations k. The irreducibility assumption was shown to imply that the latter inequalities are always strict, which entails the finite time convergence (each policy yields a unique well defined eigenvalue, these eigenvalues constitute a decreasing sequence, and there are finitely many policies).

However, the assumption that *all* stochastic matrices $P^{\sigma\tau}$ are irreducible is way too strong to guarantee that Algorithm 2 is properly posed. Indeed, to execute the algorithm, it suffices that at every iteration k the operator T^{σ_k} admits an eigenvalue and an eigenvector, which is the case in particular if for all policies σ, the graph obtained by taking the *union* of the edge sets of all the graphs associated with $P^{\sigma\tau}$ for the different choices of τ is strongly connected (see Subsection 5.2). In particular, the irreducibility assumption of Hoffman and Karp is essentially never satisfied for deterministic games, whereas the condition involving the union of the edge sets is satisfied by relevant classes of deterministic games.

It should be noted that Algorithm 2 may, in general, lead to degenerate iterations, in which $\lambda^{k+1} = \lambda^k$. As shown by an example in [ACTDG12, Sec. 6], this may lead the algorithm to cycle when the bias vector is not unique. This difficulty was solved first in the deterministic framework in [CTGG99], where it was shown that cycling can be avoided by enforcing a special choice of the bias vector, obtained by a nonlinear projection operation. This approach was then extended to the stochastic framework in [CTG06, ACTDG12]. As a special case of these results, we get that policy iteration is correct and does terminate under much milder conditions than in Theorem 7.11.

Theorem 7.12 (Cor. of [CTG06, Th. 7])**.** *Algorithm 2 terminates and is correct if for each choice of policy σ of player* MIN*, the operator T^σ has an eigenvalue and a* unique *eigenvector, up to an additive constant.*

We next show that the conditions of Theorem 7.12 are satisfied for generic payments, and conclude that nongeneric instances can still be solved by the Hoffman-Karp algorithm after an effective perturbation of the input.

7.4.2 Generic termination of policy iteration

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of a finite stochastic game with perfect information. The following assumption guarantees that Algorithm 2 is well posed for any additive perturbation of T.

Assumption 7.D. For any policy $\sigma \in \mathcal{S}_s$ of player MIN, the one-player Shapley operator T^σ is ergodic, i.e., $g + T^\sigma$ has an eigenvalue for all perturbation vectors $g \in \mathbb{R}^n$.

This assumption is much milder than the original assumption of Hoffman and Karp, requiring all the transition matrices $P^{\sigma\tau}$ to be irreducible (see Theorem 7.11). Moreover, we shall see in the next subsection that one can always transform a game (in polynomial time) by a "big M" trick in such a way that Assumption 7.D becomes satisfied.

By using the arguments of Section 7.3, we now show that under Assumption 7.D, Algorithm 2 terminates for a generic perturbation of the payments.

Theorem 7.13. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of a finite stochastic game with perfect information satisfying Assumption 7.D. Then, the space $\mathbb{R}^n$ can be covered by a polyhedral complex such that for each perturbation vector $g \in \mathbb{R}^n$ in the interior of a full-dimensional cell, Algorithm 2 terminates after a finite number of steps and gives an eigenpair of $g + T$.*

Proof. Consider the same complex $\mathcal{C}$ as in Section 7.3 and let g be a perturbation vector in the interior of a full-dimensional cell of $\mathcal{C}$. It follows from the proof of Theorem 7.8 that for any policy $\sigma \in \mathcal{S}_s$, the eigenvector of $g + T^\sigma$, which exists according to Assumption 7.D, is unique up to an additive constant. Hence, at each step k of Algorithm 2, the bias vector v^k of $g + T^{\sigma_k}$ is unique up to an additive constant. The conclusion follows from Theorem 7.12. □

We next provide an explicit perturbation g, depending on a parameter ε, for which the policy iteration algorithm applied to $g+T$ is valid. We shall see in the next subsection that ε can be instantiated with a polynomial number of bits, in such a way that the original unperturbed problem is solved.

Proposition 7.14. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of a finite stochastic game with perfect information. Then, there exists $\varepsilon_0 > 0$ such that all the perturbation vectors $g_\varepsilon := (\varepsilon, \varepsilon^2, \dots, \varepsilon^n)$ with $0 < \varepsilon < \varepsilon_0$ are in the interior of the same full-dimensional cell of the polyhedral complex of Theorem 7.13.*

Proof. The cells of the polyhedral complex $\mathcal{C}$ introduced in Theorem 7.8 that do not have a full dimension are included in an arrangement of finitely many hyperplanes. The real curve $\varepsilon \mapsto g_\varepsilon = (\varepsilon, \dots, \varepsilon^n)$ cannot cross a given hyperplane in this arrangement more than n times (otherwise, a polynomial of degree n would have strictly more than n roots). We deduce that there is a value $\varepsilon_0 > 0$ such that the restriction of the curve $\varepsilon \mapsto g_\varepsilon$ to the open interval $(0, \varepsilon_0)$ crosses no hyperplane of the arrangement. Therefore, it must stay in the interior of a full-dimensional cell of the complex $\mathcal{C}$. □

The following result is a refinement of the previous one.

Proposition 7.15. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of a finite stochastic game with perfect information. Then, there exist $\varepsilon_1 > 0$ and policies $\sigma \in \mathcal{S}_s$ and $\tau \in \mathcal{T}_s^\sigma$ such that for all $\varepsilon \in [0, \varepsilon_1]$, the upper mean payoff of $g_\varepsilon + T$ is given by*

$$\overline{\chi}(g_\varepsilon + T) = \langle m^{\sigma\tau}, g_\varepsilon + r^{\sigma\tau} \rangle \ ,$$

where $m^{\sigma\tau}$ is an invariant probability measure of the stochastic matrix $P^{\sigma\tau}$.

Proof. Recall that the upper mean payoff of $g + T$ is given by

$$\overline{\chi}(g+T) = \min_{\sigma \in \mathcal{S}_s} \overline{\chi}^\sigma(g)$$
$$\text{where} \qquad \overline{\chi}^\sigma(g) = \max\{\langle m, g + r^{\sigma\tau} \rangle \mid \tau \in \mathcal{T}_s^\sigma, \ m \in \mathcal{M}^*(P^{\sigma\tau})\} \ .$$

By construction of the polyhedral complex $\mathcal{C}$ of Theorem 7.8, in the interior of a full-dimensional cell, each piecewise affine map $\overline{\chi}^\sigma$ coincides with a unique affine map, but $g \mapsto \overline{\chi}(g+T)$ need not be affine. Hence, we can refine the complex $\mathcal{C}$ into a complex $\mathcal{C}'$ such that the latter piecewise affine map also coincides with a unique affine map on each full-dimensional cell. The exact same proof as Proposition 7.14 leads to the conclusion. □

It readily follows that solving the game with Shapley operator $g_\varepsilon + T$ for ε small enough entails a solution of the original game.

Proposition 7.16. *If T satisfies Assumption 7.D, then Algorithm 2 terminates for any input $g_\varepsilon + T$ with $\varepsilon \in (0, \varepsilon_1)$. Furthermore, any policy σ which is optimal for $g_\varepsilon + T$, meaning that $\lambda(g_\varepsilon + T) = \lambda(g_\varepsilon + T^\sigma)$, is also optimal for T, i.e., $\lambda(T) = \lambda(T^\sigma)$.* □

7.4.3 Complexity issues

In this subsection, we show that computing the upper mean payoff of a Shapley operator (a fortiori the eigenvalue if it exists) is polynomial-time reducible to the computation of the eigenvalue of a Shapley operator for which Algorithm 2 terminates. This fact is a direct consequence of Theorem 7.22 below. To do so, we shall need to give explicit bounds on the perturbation parameter ε.

We first explain how the general case can be reduced to the situation in which Assumption 7.D holds. To that purpose, let use introduce for any real number $M \geqslant 0$, the map $R_M : \mathbb{R}^n \to \mathbb{R}^n$ the ith coordinate of which is given by

$$[R_M(x)]_i = \max \left\{ x_i, \max_{j \in [n]} (-M + x_j) \right\} \ , \quad x \in \mathbb{R}^n \ .$$

Observe that R_M is a projection on the set $\{x \in \mathbb{R}^n \mid \|x\|_{\mathrm{H}} \leqslant M\}$, meaning that $R_M^2 = R_M$ and that

$$R_M(x) = x \iff \|x\|_{\mathrm{H}} \leqslant M \ .$$

Lemma 7.17. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of a perfect-information finite stochastic game, and let $M \geqslant 0$. Then, $T \circ R_M$ has an eigenvalue.*

Proof. First, note that the recession operator of $\widehat{R}_M$ is given by

$$\widehat{R}_M(x) = (\max x)\, e \ , \quad x \in \mathbb{R}^n \ ,$$

where $\max x := \max_{1 \leqslant i \leqslant n} x_i$. Second, since Assumption 4.A holds, it is readily seen that the limit (2.7) defining $\widehat{T}$ is uniform in x. Hence, we get that $\widehat{T \circ R_M} = \widehat{T} \circ \widehat{R}_M$. Thus, using the properties of recession operators, we have for any vector $x \in \mathbb{R}^n$,

$$\widehat{T \circ R_M}(x) = \widehat{T} \circ \widehat{R}_M(x) = \widehat{T}\big((\max x)\, e\big) = (\max x)\, e \ .$$

This proves that the only fixed points of $\widehat{T \circ R_M}$ are trivial. The conclusion follows from Theorem 4.6. □

The operator $T \circ R_M$ can be interpreted as the Shapley operator of a perfect-information finite stochastic game with state space $[n]$. In this game, at each step, if the current state is $i \in [n]$, player MIN start by choosing an action $a \in A_i$. Then, player MAX chooses an action $b \in B_{i,a}$ which gives rise to a transition payment $r(i, a, b)$ and a state j is chosen randomly with probability $p(j \mid i, a, b)$ and announced to the players. Finally, MAX has the possibility to choose the next state: if he pick any state k different from j, then he has an additional payment of $-M$. In other words, MAX has the option of teleporting himself to any other state, by accepting a penalty M.

Note that, since MIN has the same action space in the latter game as in the game Γ, the sets of his stationary strategies in both games are identical. Then, for a fixed policy σ of MIN, the one-player Shapley operator $(T \circ R_M)^\sigma$ is equal to $T^\sigma \circ R_M$, and we get from Lemma 7.17 the following result.

Corollary 7.18. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of a perfect-information finite stochastic game. Then, $T \circ R_M$ satisfies Assumption 7.D.*

In the modified game with Shapley operator $T \circ R_M$, player MAX makes, at each step, the final decision about the next state, provided an additional cost of M. The following result shows that if this cost is large enough, then MAX cannot do better, in the long run, than in the game Γ.

Lemma 7.19. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of a perfect-information finite stochastic game. Then, there exists a positive constant M_0 such that for any $M > M_0$, the eigenvalue of $T \circ R_M$ is equal to the upper mean payoff $\overline{\chi}(T)$.*

Proof. First, note that $R_M(x) \geqslant x$ for all $x \in \mathbb{R}^n$. Hence, by monotonicity of T, we deduce that $T \circ R_M \geqslant T$, which yields $\chi(T \circ R_M) \geqslant \chi(T)$. Since $T \circ R_M$ has an eigenvalue, denoted by $\lambda(T \circ R_M)$, then we have $\lambda(T \circ R_M) \geqslant \overline{\chi}(T)$.

By application of a Collatz-Wielandt formula (see [GG04]), we know that the eigenvalue of $T \circ R_M$ is given by

$$\lambda(T \circ R_M) = \inf\{\mu \in \mathbb{R} \mid \exists u \in \mathbb{R}^n,\ T \circ R_M(u) \leqslant \mu e + u\} \enspace .$$

We also know that T has an invariant half-line with direction $\chi(T)$. So there exists a vector $u \in \mathbb{R}^n$ such that $T(u) = u + \chi(T)$. Now let $M_0 := \|u\|_{\mathrm{H}}$. For every $M > M_0$, we have

$$T \circ R_M(u) = T(u) = u + \chi(T) \leqslant u + \overline{\chi}(T)e \enspace .$$

Hence $\lambda(T \circ R_M) \leqslant \overline{\chi}(T)$. □

We shall need a technical bound on invariant measures of stochastic matrices arising from strategies. We state it here for an arbitrary irreducible stochastic matrix.

Lemma 7.20. *Let P be a $n \times n$ irreducible stochastic matrix whose entries are rational numbers with numerators and denominators bounded by an integer D. Then, every entry of the invariant probability measure of P is a rational number whose denominator is bounded by*

$$n^{n/2} D^{n^2}$$

Proof. The invariant probability measure m of P is the unique solution of the linear system

$$\begin{cases} (I - P^\intercal)\, m = 0 \enspace , \\ e^\intercal\, m = 1 \enspace , \end{cases} \tag{7.8}$$

where I is the identity matrix. Note that one row of the subsystem $(I - P^\intercal)\, m = 0$ is redundant since we are dealing with stochastic vectors and $e^\intercal\, m = 1$. Then, by deleting this row and by multiplying every row of the latter subsystem by all the denominators of the coefficients appearing in this row, we arrive at a Cramer linear system with integer coefficients of absolute value less than D^n, and with unit coefficients on the last row. Solving this system by Cramer's rule, we obtain that the entries of m are rational numbers whose denominators divide the determinant of the system. Using Hadamard's inequality for determinants, we deduce that these denominators are bounded by

$$\big((n-1)(D^n)^2 + 1\big)^{n/2} \leqslant n^{n/2} D^{n^2} \enspace .$$

□

We just showed that the upper mean payoff of T can be recovered from the upper mean payoff of the operator $T \circ R_M$ for a suitable large M. The latter operator satisfies Assumption 7.D, and so, we can in principle apply Algorithm 2 to it. However, to do so in a way which leads in a polynomial-time transformation of the input, it is convenient to introduce the following modified Shapley operator $T_M : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ given, for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, by

$$T_M(x, y) := (T(y), R_M(x)) \enspace .$$

Note that we have

$$(T_M)^2(x, y) = \begin{pmatrix} T \circ R_M(x) \\ R_M \circ T(y) \end{pmatrix} \enspace .$$

The following immediate lemma shows that one can recover the eigenvectors and the eigenvalue of $T \circ R_M$ from those of T_M.

Lemma 7.21. *(v, w) is an eigenvector of T_M with eigenvalue λ if, and only if, v is an eigenvector of $T \circ R_M$ with eigenvalue 2λ and $w = R_M(v) - \lambda e$.* □

The operator T_M is the dynamic programming operator of a game, denoted by Γ_M, with state space $[2n]$. In each state $i \in [n]$, the actions, the payment and transition functions are the same as in Γ, except that the next state is labeled by an element of $\{n+1, \dots, 2n\}$ instead of $[n]$. Moreover, in each state $i \in \{n+1, \dots, 2n\}$, player MIN has only one possible action, while player MAX chooses the next state j among $[n]$ with a cost M if $i - j \neq n$. In particular, the policies of MIN in the two games Γ and Γ_M are in one-to-one correspondence, and to simplify the presentation, we shall use the same notation for these policies. Hence, we shall write $(T_M)^\sigma(x, y) = (T^\sigma(y), R_M(x))$ for such a policy σ.

We saw in Lemma 7.17 that the operator $T \circ R_M$ has an eigenvalue. The same is true for the operator T_M by Lemma 7.21, as well as for any operator $(T_M)^\sigma$. Thus, T_M satisfies Assumption 7.D. We know that for $\varepsilon > 0$ small enough, the perturbed operator $g_\varepsilon + T \circ R_M$ has a unique bias vector, up to an additive constant, where $g_\varepsilon = (\varepsilon, \dots, \varepsilon^n)$. This leads to considering, for $\varepsilon > 0$,

$$T_{M,\varepsilon} := (g_\varepsilon, 0) + T_M \quad .$$

Theorem 7.22. *Let Γ be a perfect-information finite stochastic game whose transition payments and probabilities are rational numbers with numerators and denominators bounded by an integer $D \geqslant 2$. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be the Shapley operator of Γ. If*

$$M > 2(2n)^{n/2} D^{2n^2+1} \quad \text{and} \quad 0 < \varepsilon < \frac{1}{n^n D^{2n(n+1)}} \ ,$$

then the upper mean payoff of T can be recovered from $T_{M,\varepsilon} = (g_\varepsilon, 0) + T_M$, in the sense that for any policy σ of player MIN such $\lambda(T_{M,\varepsilon}) = \lambda((T_{M,\varepsilon})^\sigma)$, we have $\overline{\chi}(T) = \overline{\chi}(T^\sigma)$. Furthermore, such an optimal policy can be obtained by applying Algorithm 2 with input $T_{M,\varepsilon}$.

Proof. Let $g \in [0, 1]^n$, and fix a policy σ of MIN. In the game Γ_M, consider a policy of MAX such that, when in state $i \in [n]$, he chooses some $b_i \in B_{i,\sigma(i)}$ and when in state $i \in \{n+1, \dots, 2n\}$, he chooses the next state to be some $j(i) \in [n]$. Then, the transition matrix associated with that choice of policy is the following $2n \times 2n$ block matrix:

$$\begin{pmatrix} 0 & P^{\sigma\tau} \\ Q & 0 \end{pmatrix} \ , \tag{7.9}$$

where $\tau \in \mathcal{T}_s^\sigma$ is the policy of MAX in the game Γ defined by $\tau(i) = b_i$ for each state i, and where Q is the $n \times n$ stochastic matrix whose coefficients are $Q_{ij} = 1$ for $j = j(i+n)$ and 0 otherwise.

Let $(m, m') \in \mathbb{R}^n \times \mathbb{R}^n$ be an invariant probability measure of the stochastic matrix (7.9). The vectors m and m' satisfy

$$mP^{\sigma\tau} = m' \ , \quad m'Q = m \ , \quad \langle m, e \rangle + \langle m', e \rangle = 1 \ . \tag{7.10}$$

We are interested in the eigenvalue of the perturbed one-player Shapley operator $(g, 0) + (T_M)^\sigma$. Hence, following formula (7.6), we consider the quantity

$$\gamma := \langle m, g + r^{\sigma\tau} \rangle - M \sum_{\substack{1 \leqslant i \leqslant n \\ Q_{ii} = 0}} m'_i \ .$$

If for every index i in the support of m', we have $Q_{ii} = 1$, then we deduce from the second equality in (7.10) that $m' = m$. This yields that $2m$ is an invariant probability measure of $P^{\sigma\tau}$, and that

$$\gamma = \langle m, g + r^{\sigma\tau} \rangle \leqslant \frac{1}{2}\,\overline{\chi}(g + T^{\sigma}) \enspace .$$

Note that the equality is attained in the above inequality for some policy τ and some invariant measure m.

If there is an index i in the support of m' such that $Q_{ii} = 0$, then we have

$$\gamma \leqslant \langle m, g + r^{\sigma\tau} \rangle - M\, m'_i \leqslant 1 + \max_{i,a,b} r(i,a,b) - K_1\, M \enspace ,$$

where K_1 is a positive constant such that $K_1 < m'_i$. Note that K_1 can be chosen independently of M and of the choice of policies. Then, taking $M > M_0 := (K_1)^{-1}(1 + (3/2)\|r\|_\infty)$, we obtain that

$$\gamma < \frac{1}{2}\min_{i,a,b} r(i,a,b) \leqslant \frac{1}{2}\,\overline{\chi}(g + T^{\sigma}) \enspace .$$

Thus, we have proved that, for all policies σ of MIN and for all $g \in [0,1]^n$, we have

$$\lambda\big((g,0) + (T_M)^{\sigma}\big) = \frac{1}{2}\,\overline{\chi}(g + T^{\sigma}) \enspace ,$$

as soon as $M > M_0$. In particular, the choice of the parameter ε such that $T_{M,\varepsilon}$ is a generic instance only relies on T (Proposition 7.15).

We now fix some $M > M_0$. Consider, for policies σ, σ' of MIN and τ, τ' of MAX, two distinct pairs $(m,d) \neq (m',d')$, where $m \in \mathcal{M}^*(P^{\sigma\tau})$, $m' \in \mathcal{M}^*(P^{\sigma'\tau'})$, $d := \langle m, r^{\sigma\tau} \rangle$ and $d' := \langle m', r^{\sigma'\tau'} \rangle$. We need to compare the affine maps $g \mapsto \langle m, g \rangle + d$ and $g \mapsto \langle m', g \rangle + d'$ along the curve $\varepsilon \mapsto g_\varepsilon$ with $\varepsilon \in (0,1)$.

Assume first that $d = d'$. Then $m \neq m'$ and we can select the smallest index i such that $m_i \neq m'_i$. Note that since m and m' are stochastic vectors, we necessarily have $i < n$ and we also have the existence of another index j such that $i < j \leqslant n$ and $m_j \neq m'_j$. Without loss of generality, we may assume that $m_i - m'_i > 0$. Let $K_2 \in \mathbb{R}$ be such that $0 < K_2 < m_i - m'_i$. Then, for any positive parameter $\varepsilon < K_2\, n^{-1}$, we have

$$\begin{aligned}(\langle m, g_\varepsilon \rangle + d) - (\langle m', g_\varepsilon \rangle + d') &= (m_i - m'_i)\varepsilon^i + \sum_{i<j\leqslant n} (m_j - m'_j)\varepsilon^j \\ &> K_2\, \varepsilon^i - n\varepsilon^{i+1} = \varepsilon^i\,(K_2 - n\varepsilon) > 0 \enspace .\end{aligned}$$

Assume now that $d \neq d'$, say $d > d'$, and let $K_3 \in \mathbb{R}$ be such that $0 < K_3 < d - d'$. Then, for any positive parameter $\varepsilon < K_3$, we have

$$(\langle m, g_\varepsilon \rangle + d) - (\langle m', g_\varepsilon \rangle + d') > \varepsilon^n - \varepsilon + K_3 > 0 \enspace .$$

Note that we can choose the positive constants K_2 and K_3 independently of $\sigma, \sigma', \tau, \tau', m$ and m'. Hence, the above arguments show that the set of polynomial functions $\varepsilon \mapsto \langle m, g_\varepsilon + r^{\sigma\tau} \rangle$ with $\sigma \in \mathcal{S}_s$, $\tau \in \mathcal{T}_s^{\sigma}$ and $m \in \mathcal{M}^*(P^{\sigma\tau})$, is totally ordered if ε is restricted to the interval $(0, \min\{K_2 n^{-1}, K_3\})$. Thus, the parameter ε_1 of Proposition 7.15 may be taken equal to $\min\{K_2 n^{-1}, K_3\}$.

To complete the proof, we next explain how to instantiate the constants K_1 to K_3. Let us start with K_2. It is a lower bound on the absolute values of the differences between two distinct entries (with same index) of invariant probability measures associated with the transition matrices of Γ. These differences are of the form $|p_1/q_1 - p_2/q_2| \geqslant 1/(q_1 q_2)$, where p_1, p_2, q_1, q_2 are integers. By Lemma 7.20, we know that $q_1, q_2 \leqslant n^{n/2} D^{n^2}$, and so

$$K_2 \geqslant \frac{1}{n^n D^{2n^2}} \enspace .$$

Likewise, K_3 is a lower bound on the absolute values of the differences between two distinct scalar products $\langle m, r^{\sigma\tau}\rangle$. Let $m \in \mathcal{M}(P^{\sigma\tau})$. It follows from the proof of Lemma 7.20 that the ith entry of m can be written as $m_i = p_i/q$ where p_i is an integer and $q \leqslant n^{n/2} D^{n^2}$ is an integer independent of i. Since every entry of $r^{\sigma\tau}$ has a denominator at most D, it follows that $\langle m, r^{\sigma\tau}\rangle$ is a rational number with denominator at most $n^{n/2} D^{n^2} D^n$. Therefore, the difference between two distinct values of $\langle m, r^{\sigma\tau}\rangle$ is at least $n^n D^{2n^2} D^{2n}$, and so

$$\min\{K_2 n^{-1}, K_3\} \geqslant \frac{1}{n^n D^{2n^2}} \min\left\{\frac{1}{n}, \frac{1}{D^{2n}}\right\} = \frac{1}{n^n D^{2n(n+1)}} \enspace .$$

Finally the constant K_1 is a lower bound for the positive entries of the invariant probability measures of the transition matrices (7.9) arising in the game Γ_M. A direct application of Lemma 7.20 provides the following bound:

$$K_1 \geqslant \frac{1}{(2n)^n D^{4n^2}} \enspace .$$

However, the matrices (7.9) have a particular structure, and an invariant measure m solves the linear system

$$\begin{cases} \begin{pmatrix} I & -Q^\intercal \\ -P^\intercal & I \end{pmatrix} m = 0 \\ e^\intercal\, m = 0 \end{cases}$$

where P is a transition matrix of the game Γ and Q is a stochastic matrix with entries either 0 or 1. Repeating the same arguments used to prove Lemma 7.20, we obtain that the positive entries of m have a denominator dividing an integer bounded by

$$\big((n-1)(D^n)^2 + 2\big)^{\frac{n}{2}} \, \big((D^n)^2 + 2\big)^{\frac{n-1}{2}} \, 2^{\frac{1}{2}} \leqslant (2n)^{n/2} \, D^{2n^2} \enspace .$$

Hence,

$$K_1 \geqslant \frac{1}{(2n)^{n/2} D^{2n^2}} \enspace ,$$

and

$$M_0 \leqslant (1 + (3/2)D)\; (2n)^{n/2} D^{2n^2} \enspace .$$

□

An important special case to which the method of Theorem 7.22 can be applied concerns *deterministic* mean-payoff games [GKK88, ZP96]. The input of such games can be described, as in [AGG12], by means of two matrices $A, B \in (\mathbb{Z} \cup \{-\infty\})^{m \times n}$. The corresponding Shapley operator can be written as

$$T_i(x) = \min_{j \in [m]} \big(-A_{ji} + \max_{k \in [n]} (B_{jk} + x_k) \big) \enspace , \quad \forall x \in \mathbb{R}^n \enspace , \quad \forall i \in [n] \enspace . \tag{7.11}$$

The corresponding game is played by moving a token on a graph in which n nodes, denoted by $1, \dots, n$, belong to player MIN, whereas m other nodes, denoted by $1', \dots, m'$, belong to player MAX. In state $i \in \{1,\dots,n\}$, MIN can move the token to a state $j \in \{1',\dots,m'\}$ such that $A_{ji} \neq -\infty$, receiving A_{ji}. In state j, MAX can move the token to a state $k \in \{1,\dots,n\}$ such that $B_{jk} \neq -\infty$, receiving B_{jk}. We assume that the matrix B has no identically infinite row, and that the matrix A has no identically infinite column, meaning that each player has at least one available action in each state. Then, the modified operator $T \circ R_M$ corresponds essentially to the matrix B_M in which infinite entries of B are replaced by $-M$, and the operator $g + T \circ R_M$ arises by subtracting the constant g_i to every entry in the ith column of A.

Theorem 7.23. *Let T denote the Shapley operator (7.11) of a deterministic mean-payoff game, with integer payoffs bounded in absolute value by $D \geqslant 2$. Then, for*

$$M > 4nD \quad \text{and} \quad 0 < \varepsilon < 1/n^3 \ ,$$

the policy iteration Algorithm 2 applied to the operator $g_\varepsilon + T \circ R_M$ terminates, and any optimal policy σ of player MIN, *i.e., such that $\lambda(g_\varepsilon + T \circ R_M) = \lambda(g_\varepsilon + T^\sigma \circ R_M)$, yields an optimal policy for the upper mean payoff of T, that is, $\overline{\chi}(T) = \overline{\chi}(T^\sigma)$.*

Proof. We adapt the proof of Theorem 7.22 to the case of deterministic transition matrices. In that special case, every invariant measure is uniform, with positive entries bounded below by $1/n$ if the state space has cardinality n. Hence, the constant K_1, arising as a lower bound for the positive entries of the invariant probability measures of the transition matrices in Γ_M, is such that $K_1 \geqslant 1/(2n)$, and then

$$M_0 \leqslant (2n)(1 + (3/2)D) \leqslant 4nD \ .$$

The constant K_2, which is a lower bound on the absolute values of the differences between two distinct entries of invariant measures of transition matrices in Γ, is such that $K_2 \geqslant 1/n^2$. As for K_3, it is a lower bound on the absolute values of the differences between two distinct values of $\langle m, r^{\sigma\tau} \rangle$. Since the payments are integers, every scalar product $\langle m, r^{\sigma\tau} \rangle$ is a rational number whose denominator divides the denominator of the positive entries of m. Hence, $K_3 \geqslant 1/n^2$, and the parameter ε must be lower than

$$\min\{K_2 n^{-1}, K_3\} \geqslant 1/n^3 \ .$$

□

Remark 7.24. One step in Algorithm 2 consists in computing an eigenpair (λ^k, v^k) of the reduced Shapley operator T^{σ_k} obtained by fixing the strategy σ_k of MIN. This is a simpler problem which can be solved by several known methods. We may apply, for instance, a similar policy iteration algorithm to T^{σ_k}, iterating this time in the space of policies τ of MAX. In this way, for each choice of τ, we arrive at an operator of the form $T^{\sigma_k,\tau}(x) = g + Px$, where P is a stochastic matrix which cannot in general be assumed to be irreducible. However, for one-player problems, a classical version of policy iteration, the multichain policy iteration introduced by Howard [How60] and Denardo and Fox [DF68], does allow one to determine (λ^k, v^k) without genericity conditions. Moreover, in the special case of deterministic games, the vector v^k is known to be a tropical eigenvector and λ^k a tropical eigenvalue. The tropical eigenpair can be computed by direct combinatorial algorithms, see e.g. the discussion in [CTGG99].

Remark 7.25. Theorem 7.23 should be compared with the other known perturbation scheme, relying on vanishing discount. The latter method requires the computation of a fixed point of the operator $x \mapsto T(\alpha x)$ for $0 < \alpha < 1$ sufficiently close to one. It is known that, for deterministic mean-payoff games, if the discount factor α is chosen so that

$$\alpha > 1 - \frac{1}{4(n+m)^3 D} \ ,$$

where D denotes the maximal absolute value of a finite entry A_{ij} or B_{ij}, then, the solution of the mean-payoff problem can be derived from the solution of the discounted problem, see [ZP96, Sec. 5]. The latter can be obtained by policy iteration (which terminates without any nondegeneracy conditions in the discounted case). The present perturbation scheme requires shorter rational numbers. Note in particular that applying Algorithm 2 to the map $g_\varepsilon + T$ requires solving a linear systems in which the matrix is independent of ε and well conditioned, whereas vanishing discount requires the inversion of a matrix which becomes singular as $\alpha \to 1$. If the vanishing discount approach is interpreted in exact arithmetics, it leads to a blow up of the bitsize of the intermediate data as $\alpha \to 1$, whereas if this is interpreted in floating point arithmetics, it may lead to numerical instabilities or overflows.

Remark 7.26. The present approach allows one to compute the upper mean payoff, i.e., the maximum of the mean payoff over all initial states. This leads to no loss of expressivity since it follows from known reductions that this problem is polynomial-time equivalent to solving a mean-payoff game in which the initial state is fixed (combine [AGS16, Cor. C.3] with the reductions in [AM09]). An alternative route to compute the mean payoff of any given initial state, avoiding the use of such reductions, would be to extend the present perturbation scheme to the "multichain" version of policy iteration, discussed in [CTG06, ACTDG12].

CHAPTER 8

Accretive operator approach to generic uniqueness of the bias vector

The results persented in this chapter have been announced, in a milder version, in the MTNS conference proceedings [Hoc16].

8.1 Generic uniqueness of the fixed point of nonexpansive maps

8.1.1 Uniqueness and continuity of the fixed-point map

Let $(\mathcal{X}, \|\cdot\|)$ be a real vector space of finite dimension, endowed with a given norm. In this section, we fix an operator $T : \mathcal{X} \to \mathcal{X}$, nonexpansive with respect to the norm of $\mathcal{X}$, and we assume that there exists an open subset $\mathcal{V} \subset \mathcal{X}$ such that $g + T$ has at least one fixed point for all $g \in \mathcal{V}$.

Let us introduce the set-valued map $\mathrm{FP} : \mathcal{V} \rightrightarrows \mathcal{X}$ defined by

$$\mathrm{FP}(g) := \{x \in \mathcal{X} \mid x - T(x) = g\} \ , \quad g \in \mathcal{V} \ .$$

Put in words, the mapping FP sends each vector $g \in \mathcal{V}$ to the set of fixed points of $g + T$. Observe that the inverse map of FP is

$$\mathrm{FP}^{-1} = \mathrm{Id} - T \ ,$$

so that it is coaccretive by Lemma 4.19. Since we have assumed that $\mathrm{dom}(\mathrm{FP}) = \mathcal{V}$, then we get from the local boundedness of coaccretive maps (Proposition 4.13) the following.

Lemma 8.1. *Let $(\mathcal{X}, \|\cdot\|)$ be a finite-dimensional real vector space, and let $T : \mathcal{X} \to \mathcal{X}$ be a nonexpansive map such that $g + T$ has a fixed point for all vectors g in an open subset $\mathcal{V} \subset \mathcal{X}$. Then, the set-valued map $\mathrm{FP} : \mathcal{V} \rightrightarrows \mathcal{X}$ has compact values and is upper semicontinuous.*

Proof. It readily follows from Proposition 4.13 that $\mathrm{FP}(g)$ is bounded for any $g \in \mathcal{V}$. Furthermore, since $g + T$ is continuous, $\mathrm{FP}(g)$ is closed, hence compact.

It is now possible to characterize sequentially the upper-semicontinuity of FP. Let $(g_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{V}$ converging to some vector $g \in \mathcal{V}$ and for all integers k, let $x_k \in \mathrm{FP}(g_k)$.

We know from Proposition 4.13 that FP is locally bounded at g. Hence the sequence $(x_k)_{k\in\mathbb{N}}$ is bounded, and therefore has an accumulation point x. Since $x_k - T(x_k) = g_k$ for all k, the continuity of T yields that $x - T(x) = g$, that is, $x \in \mathrm{FP}(g)$. This proves that FP is upper semicontinuous at g. □

We shall need the following technical result, which is a variant of the Hahn-Banach theorem.

Lemma 8.2. *Let J be the duality mapping on $(\mathcal{X}, \|\cdot\|)$, and let x be any vector in $\mathcal{X}$. Then,*

$$x \neq 0 \quad \Longleftrightarrow \quad \exists y \in \mathcal{X} \setminus \{0\} \ , \quad \forall x^* \in J(x) \ , \quad \langle y, x^* \rangle > 0 \ .$$

Proof. Let $x \in \mathcal{X} \setminus \{0\}$. We know that $J(x)$ is a compact convex subset of $\mathcal{X}^*$. Furthermore, $0 \notin J(x)$. Hence, according to the Hahn-Banach theorem, there exists an affine hyperplane of $\mathcal{X}^*$ strongly separating the two compact convex subsets $J(x)$ and $\{0\}$, i.e., there exists some vector $y \in \mathcal{X} \setminus \{0\}$ and a constant $\varepsilon > 0$ such that for all $x^* \in J(x)$ we have $\langle y, x^* \rangle \geqslant \varepsilon \geqslant 0$.

Conversely, if $x = 0$, then $J(x) = \{0\}$ and so, for all $y \in \mathcal{X}$ we have $\langle y, x^* \rangle = 0$ with $x^* = 0 \in J(x)$. □

We now state the main result of this section.

Theorem 8.3. *Let $(\mathcal{X}, \|\cdot\|)$ be a finite-dimensional real vector space, and let $T : \mathcal{X} \to \mathcal{X}$ be a non-expansive map such that $g + T$ has a fixed point for all vectors g in an open subset $\mathcal{V} \subset \mathcal{X}$. Then, the set-valued map* $\mathrm{FP} : \mathcal{V} \rightrightarrows \mathcal{X}$ *is continuous at point $g \in \mathcal{X}$ if, and only if,* $\mathrm{FP}(g)$ *is a singleton.*

Proof. Suppose first that the map FP is single-valued at g. We have shown in Lemma 8.1 that FP is upper semicontinuous at g. It readily follows from the definition (Subsection 4.3.1) that FP is also lower semicontinuous at g, since it is single-valued. Hence, FP is continuous at g.

Conversely, suppose that FP is continuous at point $g \in \mathcal{V}$ and let x and x' be two vectors in $\mathrm{FP}(g)$, i.e., two fixed points of $g + T$. Choose $y \in \mathcal{X} \setminus \{0\}$, and for every positive integer k, let $g_k := g - k^{-1}y$. We may assume, without loss of generality, that g_k is in $\mathcal{V}$ for all k. Since FP is continuous at g, hence lower semicontinuous, there exists a sequence of elements $x_k \in \mathrm{FP}(g_k)$ converging to x. Furthermore, we know from Lemma 4.19 that the map $\mathrm{FP} = (\mathrm{Id} - T)^{-1}$ is coaccretive. Hence, for every integer $k > 0$, there is a point $x_k^* \in J(x_k - x')$ such that

$$\langle g_k - g, x_k^* \rangle \geqslant 0 \ ,$$

which yields

$$\langle y, x_k^* \rangle \leqslant 0 \ .$$

Since the duality mapping J is upper semicontinuous (Proposition 4.8) and with compact values, then the sequence of elements $x_k^* \in J(x_k - x')$ has an accumulation point $x^* \in J(x - x')$, which satisfies

$$\langle y, x^* \rangle \leqslant 0 \ .$$

Thus, we have proved that for any point $y \in \mathcal{X} \setminus \{0\}$, there exists an element $x^* \in J(x - x')$ such that $\langle y, x^* \rangle \leqslant 0$. We deduce from Lemma 8.2 that $x - x' = 0$, and consequently that $\mathrm{FP}(g)$ is a singleton. □

8.1.2 Generic continuity

Under mild conditions, the continuity of any upper semicontinuous set-valued map A is a generic property. By *generic*, we mean that the set of points where the map A is continuous is a residual. Recall that a *residual* of any space $\mathcal{Y}$ is a countable intersection of dense open subsets of $\mathcal{Y}$. In particular, if $\mathcal{Y}$ is a Baire space, then a residual is dense.

Theorem 8.4 (see [AF09, Th. 1.4.13])**.** *Let A be a set-valued map from a complete metric space $\mathcal{X}$ to a complete separable metric space $\mathcal{Y}$. If A is upper semicontinuous on $\mathcal{X}$, then it is continuous on a residual of $\mathcal{X}$.*

Note that Theorem 8.4 states that the set of points where an upper semicontinous set-valued map is continuous (only) contains a residual. We can in fact be more specific about the characterization of elements where the fixed-point map FP is single-valued.

Theorem 8.5. *Let $(\mathcal{X}, \|\cdot\|)$ be a finite-dimensional real vector space, and let $T : \mathcal{X} \to \mathcal{X}$ be a nonexpansive map. Let $\mathcal{U}$ be an open subset of $\mathcal{X}$ such that $g + T$ has a fixed point for all vectors g in a neighborhood of $\overline{\mathcal{U}}$. Then, the set of vectors $g \in \overline{\mathcal{U}}$ for which $g + T$ has a unique fixed point is a residual of $\overline{\mathcal{U}}$. In particular, it is dense in $\overline{\mathcal{U}}$.*

Proof. According to Lemma 8.1, the set-valued map FP is upper semicontinuous on a neighborhood of $\overline{\mathcal{U}}$. Hence, its restriction to $\overline{\mathcal{U}}$ (which is also upper semicontinuous) is continuous on a dense subset (Theorem 8.4). Consider the intersection of this dense subset with $\mathcal{U}$, which is still dense in $\overline{\mathcal{U}}$. According to Theorem 8.3, at any point g in the latter subset, $g+T$ has a unique fixed point. So there is a dense subset of $\overline{\mathcal{U}}$ such that for each point g in this subset, $g + T$ has a unique fixed point. It remains to show that the set of points in $\overline{\mathcal{U}}$ for which FP is a singleton is a countable intersection of open subsets.

There is no loss of generality in assuming that $0 \in \mathcal{U}$. Let $(y_i)_{i\geqslant 1}$ be a countable dense subset of $\mathcal{U}$. For all positive integers i and k, let

$$\mathcal{F}_{i,k} := \{g \in \overline{\mathcal{U}} \mid \exists x, x' \in \mathrm{FP}(g),\ \ \exists x^* \in J(x - x'),\ \ \langle y_i, x^*\rangle \geqslant 1/k\} \ .$$

We show that FP is single-valued at g if, and only if, $g \notin \bigcup_{i,k\in\mathbb{N}} \mathcal{F}_{i,k}$. First, if $\mathrm{FP}(g)$ is a singleton, then g cannot be in any set $\mathcal{F}_{i,k}$. Conversely, if it is not a singleton, we can find two distinct fixed points of $g + T$, x, x', and according to Lemma 8.2, there exists some vector $y \in \mathcal{X} \setminus \{0\}$ such that $\langle y, x^*\rangle > 0$ for all $x^* \in J(x - x')$. Since $\mathcal{U}$ is an open set containing 0, we may assume, up to a scaling, that y is in $\mathcal{U}$. Therefore, y can be approximated by some point y_i such that the inequality $\langle y_i, x^*\rangle > 0$ still holds for all $x^* \in J(x - x')$. Hence $g \in \mathcal{F}_{i,k}$ for some integer k large enough.

We next show that all the sets $\mathcal{F}_{i,k}$ are closed, which will prove that FP is single-valued on a residual of $\overline{\mathcal{U}}$. Fix the integers $i, k \geqslant 1$ and let (g_ℓ) be a sequence in $F_{i,k}$ converging to some vector g. For every integer ℓ, there exist $x_\ell, x'_\ell \in \mathrm{FP}(g_\ell)$ and $x^*_\ell \in J(x_\ell - x'_\ell)$ such that $\langle y_i, x^*_\ell\rangle \geqslant 1/k$. Since the map FP is locally bounded at g (Proposition 4.13), we may suppose, up to an extraction, that the sequences (x_ℓ) and (x'_ℓ) converge to some points x and x', respectively. By continuity of T, it is readily seen that $x, x' \in \mathrm{FP}(g)$. Furthermore, since the duality mapping J has compact values and is upper semicontinuous, the sequence (x^*_ℓ) has an accumulation point x^* in $J(x - x')$, which satisfies $\langle y_i, x^*\rangle \geqslant 1/k$. Thus $g \in F_{i,k}$. □

8.2 Application to nonlinear Perron-Frobenius theory

We apply the main result of the previous section (Theorem 8.5) to the case of Shapley operators arising from ergodic stochastic games.

Theorem 8.6. *Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an ergodic Shapley operator. Then, the set of vectors g for which $g + T$ has a unique bias vector, up to an additive constant, is a residual of $\mathbb{R}^n$.*

Proof. We apply Theorem 8.5 to the quotiented self-map $[T]$ of the additive projective space $\mathbb{TP}^n$, endowed with the norm q_{H}. Thus, there exists a residual of $\mathbb{TP}^n$, denoted by $\tilde{\mathcal{R}}$, characterizing the set of equivalence classes $[g]$ for which the quotiented map $[g] + [T] = [g + T]$ has a unique fixed point. Hence, the operator $g + T$ has a unique bias vector up to an additive constant if, and only if, the equivalence class of g is in this residual.

Let $\mathcal{R}$ be the preimage in $\mathbb{R}^n$ of $\tilde{\mathcal{R}}$ by the canonical map $\mathbb{R}^n \to \mathbb{TP}^n$. The latter quotient map is continuous, since the preimage of any open ball, that is, any set $\{x \in \mathbb{R}^n \mid \|x - g\|_{\mathrm{H}} < \rho\}$ with $\rho > 0$ and $g \in \mathbb{R}^n$, is open. Hence, it follows that $\mathcal{R}$ is also a countable intersection of open subsets. Moreover, for any vector $g \in \mathbb{R}^n$ and any real number $\varepsilon > 0$, there exists a vector $h \in \mathcal{R}$ such that $q_{\mathrm{H}}([g] - [h]) = \|g - h\|_{\mathrm{H}} < \varepsilon$, since the residual $\tilde{\mathcal{R}}$ is dense in $\mathbb{TP}^n$. The vector h is defined up to an additive constant, so we may assume that

$$\min_{i \in [n]} (g_i - h_i) = 0 \ ,$$

in which case $\|g - h\|_{\mathrm{H}}$ is equal to the sup-norm of $g - h$. Therefore, for any vector in $\mathbb{R}^n$ we can find a vector in $\mathcal{R}$ arbitrarily close. This shows that $\mathcal{R}$ is dense in $\mathbb{R}^n$. □

CHAPTER 9

Conclusion and perspectives

We now briefly summarize our main contributions and point out some open problems.

In Chapters 4 to 6, we have introduced a notion of ergodicity for zero-sum stochastic games with finite state space, which recover the finite Markov chain case. We have also given several characterizations for this property to hold, in particular in terms of hypergraph reachability. To that purpose, we have followed the operator approach by applying techniques from nonlinear Perron-Frobenius theory to the study of Shapley operators. The perspectives of this work follow naturally two paths, one in relation with game-theoretic aspects, and the other with nonlinear Perron-Frobenius theory.

Regarding nonlinear Perron-Frobenius theory, we have given a condition that guarantees the existence of a fixed point for any additive perturbation of a nonexpansive map T (Corollary 4.20) for any norm in a finite-dimensional vector space. This condition, established in the additive framework, requires that all the spaces $\{x \mid \|x - T(x)\| \leqslant \alpha\}$ are bounded. A natural question is to know if a similar result holds in the multiplicative framework. More precisely, let us consider a symmetric cone K of a Euclidean space, i.e., a self-dual open convex cone, the automorphism group G of which acts transitively on it. Let $f : K \to K$ be an order-preserving homogeneous map, such that the self-map of K sending x to $g(f(x))$ has a fixed point for all automorphisms $g \in G$. Does this imply that all the spaces $\mathcal{D}_\alpha(f) := \{x \in K \mid \mathrm{d_H}(x, f(x)) \leqslant \alpha\}$ are bounded in Hilbert's projective metric $\mathrm{d_H}$ arising from K? In particular, is this true when K is the cone of positive semi-definite matrices? Another problem, following the results in Chapter 6, is to find a combinatorial characterization, similar to the one in Theorem 6.2, of the boundedness of all spaces $\mathcal{D}_\alpha(f)$ for any cone K. When the cone K is simplicial, this problem should reduce to the case of the standard nonnegative cone of $\mathbb{R}^n$ treated in this thesis. When K is polyhedral but not simplicial, we expect that the characterization involves the asymptotic behavior of f along the faces of K.

As for game theory problems, it would be interesting to extend the characterization of ergodicity to stochastic games with infinite state space. It would allow in particular to broaden the latter notion to other classes of zero-sum repeated games, including games with incomplete information. Indeed, a wide range of repeated games have a recursive structure, expressing itself in a Shapley operator, see [Sor03] and references therein. It is then possible to reduce these games to stochastic games, the state space of which is a probability space over the set of un-known parameters.

A first step to solve the above problem could be to generalize the surjectivity conditions for accretive maps (Theorem 4.16) to the case of infinite-dimensional Banach spaces. However, this generalization would most likely require strong topological assumptions (reflexive Banach space, fixed-point property) which might not be compatible with our framework (say space of continuous or bounded functions on a compact set, endowed with Hilbert's seminorm). Another idea would be to start directly from the formulation of ergodicity in terms of dominions and discretize the infinite state space in order to approximate the original game with a simpler one with finite state space.

With continuous time and space (i.e., differential games), the problem of the asymptotic behavior of the value in finite horizon, known as ergodic problem for Hamilton-Jacobi equations, has attracted much attention lately, see [Bar09, AB10]. In the deterministic framework, the value $u(x,t)$ in horizon t with initial state x is the viscosity solution of an Hamilton-Jacobi PDE

$$\partial_t u + H(x, \partial_x u) = 0 \ ,$$

for some Hamiltonian H. An important question is then to understand when $u(x,t)/t$ converges as $t \to +\infty$ to some constant λ. Note that if this property holds, λ is the unique constant c for which the Hamilton-Jacobi PDE

$$H(x, \partial_x v) = -c$$

has a viscosity solution v. Hence, the latter equation replaces the ergodic equation (2.5) when time and space are continuous. In the one-player case, i.e., when $(x,p) \mapsto H(x,p)$ is convex in the adjoint variable p, the study of the ergodic problem is treated in the framework of weak KAM theory. However, in the two-player case (nonconvex Hamiltonian), weak KAM methods no longer apply and the problem is less well understood [Car10]. Although the results of this thesis do not extend easily to the PDE case, we hope they could provide a good inspiration to tackle the latter issues. In particular, an operator approach is possible by considering the evolution semigroup associated with the game.

In Chapters 7 and 8, we have used tools of nonlinear Perron-Frobenius theory to establish the uniqueness, up to an additive constant, of the bias vector of an ergodic Shapley operator for a generic perturbation. We have also shown more generally that this generic property is true locally, as soon as the ergodic equation is solvable for all additive perturbations in any open subset. These results make a contribution to the more general problem of understanding the structure of the set of bias vectors of Shapley operators.

In the case of payment-free Shapley operators, the results in Chapter 5 give partial answers to the latter problem, which seems hard to tackle in full generality. The case of Shapley operators with perfect information and deterministic transitions is already interesting. When the operator is Boolean monotone (corresponding to payment-free Shapley operator with perfect information and deterministic transitions), it is possible to show that the set of bias vectors (which coincides with the set of fixed points) is uniquely determined by the Boolean fixed points (i.e., the coordinates of which are either 0 or 1). More generally, in the deterministic perfect-information framework, the set of bias vectors is a polyhedral complex, intersection between two tropical convex sets, one max-plus and the other min-plus, but it remains an open question to find a suitable description of this space.

Bibliography

[AB06] C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis*. Springer, Berlin, third edition, 2006.

[AB09] S. Arora and B. Barak. *Computational complexity*. Cambridge University Press, Cambridge, 2009.

[AB10] O. Alvarez and M. Bardi. Ergodicity, stabilization, and singular perturbations for Bellman-Isaacs equations. *Mem. Amer. Math. Soc.*, 204(960):vi+77, 2010.

[AB15] V. Anantharam and V. S. Borkar. A variational formula for risk-sensitive reward. arXiv:1501.00676, 2015.

[ABG13] M. Akian, R. B. Bapat, and S. Gaubert. Max-plus algebras. In Leslie Hogben, editor, *Handbook of Linear Algebra, Second Edition*, Discrete Mathematics and its Applications, chapter 35. Chapman & Hall/CRC, Boca Raton, FL, 2013.

[ACTDG12] M. Akian, J. Cochet-Terrasson, S. Detournay, and S. Gaubert. Policy iteration algorithm for zero-sum multichain stochastic games with mean payoff and perfect information. arXiv:1208.0446, 2012.

[ADEH99] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999.

[AF09] J.-P. Aubin and H. Frankowska. *Set-valued analysis*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2009. Reprint of the 1990 edition.

[AG03] M. Akian and S. Gaubert. Spectral theorem for convex monotone homogeneous maps, and ergodic control. *Nonlinear Anal.*, 52(2):637–679, 2003.

[AGG12] M. Akian, S. Gaubert, and A. Guterman. Tropical polyhedra are equivalent to mean payoff games. *Internat. J. Algebra Comput.*, 22(1):1250001, 43, 2012.

[AGH14a] M. Akian, S. Gaubert, and A. Hochart. Fixed point sets of payment-free Shapley operators and structural properties of mean payoff games. In *21st International Symposium on Mathematical Theory of Networks and Systems*, pages 1438–1441, Groningen, The Netherlands, July 2014.

[AGH14b] M. Akian, S. Gaubert, and A. Hochart. Generic uniqueness of the bias vector of mean payoff zero-sum games. In *53rd IEEE Conference on Decision and Control*, pages 1581–1587, Los Angeles, California, USA, December 2014. arXiv:1411.1211.

[AGH15a] M. Akian, S. Gaubert, and A. Hochart. Ergodicity conditions for zero-sum games. *Discrete Contin. Dyn. Syst.*, 35(9):3901–3931, 2015.

[AGH15b] M. Akian, S. Gaubert, and A. Hochart. Hypergraph conditions for the solvability of the ergodic equation for zero-sum games. In *54th IEEE Conference on Decision and Control*, pages 5845–5850, Osaka, Japan, December 2015. arXiv:1510.05396.

[AGH16] M. Akian, S. Gaubert, and A. Hochart. Minimax representation of nonexpansive functions and application to zero-sum recursive games. arXiv:1605.04518, 2016.

[AGL08] M. Akian, S. Gaubert, and A. Lakhoua. The max-plus finite element method for solving deterministic optimal control problems: basic properties and convergence analysis. *SIAM J. Control Optim.*, 47(2):817–848, 2008.

[AGN16] M. Akian, S. Gaubert, and R. Nussbaum. Uniqueness of the fixed point of nonexpansive semidifferentiable maps. *Trans. Amer. Math. Soc.*, 368(2):1271–1320, 2016.

[AGS16] X. Allamigeon, S. Gaubert, and M. Skorma. Solving generic nonarchimedean semidefinite programs using stochastic game algorithms. arXiv:1603.06916, 2016.

[AGW09] M. Akian, S. Gaubert, and C. Walsh. The max-plus Martin boundary. *Doc. Math.*, 14:195–240, 2009.

[All14] X. Allamigeon. On the complexity of strongly connected components in directed hypergraphs. *Algorithmica*, 69(2):335–369, 2014.

[AM95] R. J. Aumann and M. B. Maschler. *Repeated games with incomplete information*. MIT Press, Cambridge, MA, 1995. With the collaboration of R. E. Stearns.

[AM09] D. Andersson and P. B. Miltersen. The complexity of solving stochastic games on graphs. In *Algorithms and computation*, volume 5878 of *Lecture Notes in Comput. Sci.*, pages 112–121. Springer, Berlin, 2009.

[Ari97] M. Arisawa. Ergodic problem for the Hamilton-Jacobi-Bellman equation. I. Existence of the ergodic attractor. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 14(4):415–438, 1997.

[Ari98] M. Arisawa. Ergodic problem for the Hamilton-Jacobi-Bellman equation. II. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 15(1):1–24, 1998.

[Asp67] E. Asplund. Positivity of duality mappings. *Bull. Amer. Math. Soc.*, 73:200–203, 1967.

[AT07] C. D. Aliprantis and R. Tourky. *Cones and duality*, volume 84 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2007.

[Bar09] M. Bardi. On differential games with long-time-average cost. In *Advances in dynamic games and their applications*, volume 10 of *Ann. Internat. Soc. Dynam. Games*, pages 3–18. Birkhäuser Boston, Inc., Boston, MA, 2009.

[Bat73] J. Bather. Optimal decision procedures for finite Markov chains. II. Communicating systems. *Advances in Appl. Probability*, 5:521–540, 1973.

[BC11] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2011.

[BCOQ92] F. L. Baccelli, G. Cohen, G. J. Olsder, and J.-P. Quadrat. *Synchronization and linearity*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Ltd., Chichester, 1992.

[BdF66] F. E. Browder and D. G. de Figueiredo. J-monotone nonlinear operators in Banach spaces. *Nederl. Akad. Wetensch. Proc. Ser. A 69=Indag. Math.*, 28:412–420, 1966.

[BGV15] J. Bolte, S. Gaubert, and G. Vigeral. Definable zero-sum stochastic games. *Math. Oper. Res.*, 40(1):171–191, 2015.

[Bir79] G. Birkhoff. *Lattice theory*, volume 25 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, R.I., third edition, 1979.

[BJ72] T. S. Blyth and M. F. Janowitz. *Residuation theory*. Pergamon Press, Oxford-New York-Toronto, Ont., 1972. International Series of Monographs in Pure and Applied Mathematics, Vol. 102.

[BK76] T. Bewley and E. Kohlberg. The asymptotic theory of stochastic games. *Math. Oper. Res.*, 1(3):197–208, 1976.

[BL62] A. Beurling and A. E. Livingston. A theorem on duality mappings in Banach spaces. *Ark. Mat.*, 4:405–411 (1962), 1962.

[Bla65] D. Blackwell. Discounted dynamic programming. *Ann. Math. Statist.*, 36:226–235, 1965.

[BP94] A. Berman and R. J. Plemmons. *Nonnegative matrices in the mathematical sciences*, volume 9 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. Revised reprint of the 1979 original.

[BR14] M. Bourque and T. E. S. Raghavan. Policy improvement for perfect information additive reward and additive transition stochastic games with discounted and average payoffs. *J. Dyn. Games*, 1(3):347–361, 2014.

[Bré73] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973.

[Bro65] F. E. Browder. Multi-valued monotone nonlinear mappings and duality mappings in Banach spaces. *Trans. Amer. Math. Soc.*, 118:338–351, 1965.

[Bro66] F. E. Browder. *Problèmes nonlinéaires*. Séminaire de Mathématiques Supérieures, No. 15 (Été, 1965). Les Presses de l'Université de Montréal, Montreal, Que., 1966.

[Bro67a] F. E. Browder. Nonlinear accretive operators in Banach spaces. *Bull. Amer. Math. Soc.*, 73:470–476, 1967.

[Bro67b] F. E. Browder. Nonlinear mappings of nonexpansive and accretive type in Banach spaces. *Bull. Amer. Math. Soc.*, 73:875–882, 1967.

[Bro68] F. E. Browder. Nonlinear monotone and accretive operators in Banach spaces. *Proc. Nat. Acad. Sci. U.S.A.*, 61:388–393, 1968.

[Bro76] F. E. Browder. Nonlinear operators and nonlinear equations of evolution in Banach spaces. In *Nonlinear functional analysis (Proc. Sympos. Pure Math., Vol. XVIII, Part 2, Chicago, Ill., 1968)*, pages 1–308. Amer. Math. Soc., Providence, R. I., 1976.

[Bru73] R. E. Bruck, Jr. Properties of fixed-point sets of nonexpansive mappings in Banach spaces. *Trans. Amer. Math. Soc.*, 179:251–262, 1973.

[Car10] P. Cardaliaguet. Ergodicity of Hamilton-Jacobi equations with a noncoercive nonconvex Hamiltonian in $\mathbb{R}^2/\mathbb{Z}^2$. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 27(3):837–856, 2010.

[CCHH05] R. Cavazos-Cadena and D. Hernández-Hernández. A characterization of the optimal risk-sensitive average cost in finite controlled Markov chains. *Ann. Appl. Probab.*, 15(1A):175–212, 2005.

[CCHH09] R. Cavazos-Cadena and D. Hernández-Hernández. Necessary and sufficient conditions for a solution to the risk-sensitive Poisson equation on a finite state space. *Systems Control Lett.*, 58(4):254–258, 2009.

[CCHH10] R. Cavazos-Cadena and D. Hernández-Hernández. Poisson equations associated with a homogeneous and monotone function: necessary and sufficient conditions for a solution in a weakly convex case. *Nonlinear Anal.*, 72(7-8):3303–3313, 2010.

[CG79] R. Cuninghame-Green. *Minimax algebra*, volume 166 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin-New York, 1979.

[CL71] M. G. Crandall and T. M. Liggett. Generation of semi-groups of nonlinear transformations on general Banach spaces. *Amer. J. Math.*, 93:265–298, 1971.

[Con92] A. Condon. The complexity of stochastic games. *Inform. and Comput.*, 96(2):203–224, 1992.

[CP72] M. G. Crandall and A. Pazy. Nonlinear evolution equations in Banach spaces. *Israel J. Math.*, 11:57–94, 1972.

[CQZ13] K. Chang, L. Qi, and T. Zhang. A survey on the spectral theory of nonnegative tensors. *Numer. Linear Algebra Appl.*, 20(6):891–912, 2013.

[CT80] M. G. Crandall and L. Tartar. Some relations between nonexpansive and order preserving mappings. *Proc. Amer. Math. Soc.*, 78(3):385–390, 1980.

[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.

[CTGG99] J. Cochet-Terrasson, S. Gaubert, and J. Gunawardena. A constructive fixed point theorem for min-max functions. *Dynam. Stability Systems*, 14(4):407–433, 1999.

[Del02] F. Delbaen. Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, pages 1–37. Springer, Berlin, 2002.

[DF68] E. V. Denardo and B. L. Fox. Multichain Markov renewal programs. *SIAM J. Appl. Math.*, 16:468–487, 1968.

[DLRS10] J. A. De Loera, J. Rambau, and F. Santos. *Triangulations*, volume 25 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, 2010.

[DMLR04] J. Dutta, J. E. Martínez-Legaz, and A. M. Rubinov. Monotonic analysis over cones. I. *Optimization*, 53(2):129–146, 2004.

[DMLR08] J. Dutta, J. E. Martínez-Legaz, and A. M. Rubinov. Monotonic analysis over cones. III. *J. Convex Anal.*, 15(3):561–579, 2008.

[DS88] N. Dunford and J. T. Schwartz. *Linear operators. Part I.* Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. General theory, With the assistance of W. G. Bade and R. G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[EM79] A. Ehrenfeucht and J. Mycielski. Positional strategies for mean payoff games. *Internat. J. Game Theory*, 8(2):109–113, 1979.

[ET99] I. Ekeland and R. Témam. *Convex analysis and variational problems*, volume 28 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, english edition, 1999. Translated from the French.

[Eva84] L. C. Evans. Some min-max methods for the Hamilton-Jacobi equation. *Indiana Univ. Math. J.*, 33(1):31–50, 1984.

[Eve57] H. Everett. Recursive games. In *Contributions to the theory of games, vol. 3*, Annals of Mathematics Studies, no. 39, pages 47–78. Princeton University Press, Princeton, N. J., 1957.

[Fat08] A. Fathi. Weak KAM theorem in Lagrangian dynamics. Tenth preliminary version, available online, 2008.

[FGH13] S. Friedland, S. Gaubert, and L. Han. Perron-Frobenius theorem for nonnegative multilinear forms and extensions. *Linear Algebra Appl.*, 438(2):738–749, 2013.

[FHH97] W. H. Fleming and D. Hernández-Hernández. Risk-sensitive control of finite state machines on an infinite horizon. I. *SIAM J. Control Optim.*, 35(5):1790–1810, 1997.

[FHH99] W. H. Fleming and D. Hernández-Hernández. Risk-sensitive control of finite state machines on an infinite horizon. II. *SIAM J. Control Optim.*, 37(4):1048–1069 (electronic), 1999.

[FHK72] P. M. Fitzpatrick, P. Hess, and T. Kato. Local boundedness of monotone-type operators. *Proc. Japan Acad.*, 48:275–277, 1972.

[FP96] K. Fukuda and A. Prodon. Double description method revisited. In *Combinatorics and computer science (Brest, 1995)*, volume 1120 of *Lecture Notes in Comput. Sci.*, pages 91–111. Springer, Berlin, 1996.

[FR13] A. Figalli and L. Rifford. Aubry sets, Hamilton-Jacobi equations, and the Mañé conjecture. In *Geometric analysis, mathematical relativity, and nonlinear partial differential equations*, volume 599 of *Contemp. Math.*, pages 83–104. Amer. Math. Soc., Providence, RI, 2013.

[FS02] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance Stoch.*, 6(4):429–447, 2002.

[FS11] H. Föllmer and A. Schied. *Stochastic finance*. Walter de Gruyter & Co., Berlin, extended edition, 2011.

[FV97] J. Filar and K. Vrieze. *Competitive Markov decision processes*. Springer-Verlag, New York, 1997.

[GG04] S. Gaubert and J. Gunawardena. The Perron-Frobenius theorem for homogeneous, monotone functions. *Trans. Amer. Math. Soc.*, 356(12):4931–4950 (electronic), 2004.

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

[GK95] J. Gunawardena and M. Keane. On the existence of cycle times for some nonexpansive maps. Technical report, Technical Report HPL-BRIMS-95-003, Hewlett-Packard Labs, 1995.

[GKK88] V. A. Gurvich, A. V. Karzanov, and L. G. Khachiyan. Cyclic games and finding minimax mean cycles in digraphs. *Zh. Vychisl. Mat. i Mat. Fiz.*, 28(9):1407–1417, 1439, 1988.

[GLNP93] G. Gallo, G. Longo, S. Nguyen, and S. Pallottino. Directed hypergraphs and applications. *Discrete Appl. Math.*, 42(2-3):177–201, 1993.

[GM77] M. Gondran and M. Minoux. Valeurs propres et vecteurs propres dans les dioïdes et leur interprétation en théorie des graphes. *Bull. Direction Études Recherches Sér. C Math. Informat.*, (2):i, 25–41, 1977.

[Gun03] J. Gunawardena. From max-plus algebra to nonexpansive mappings: a nonlinear theory for discrete event systems. *Theoret. Comput. Sci.*, 293(1):141–167, 2003.

[GV12] S. Gaubert and G. Vigeral. A maximin characterisation of the escape rate of nonexpansive mappings in metrically convex spaces. *Math. Proc. Cambridge Philos. Soc.*, 152(2):341–363, 2012.

[HHM96] D. Hernández-Hernández and S. I. Marcus. Risk sensitive control of Markov processes in countable state space. *Systems Control Lett.*, 29(3):147–155, 1996.

[HK66] A. J. Hoffman and R. M. Karp. On nonterminating stochastic games. *Management Sci.*, 12:359–370, 1966.

[HLL96] O. Hernández-Lerma and J. B. Lasserre. *Discrete-time Markov control processes*, volume 30 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1996.

[HLL99] O. Hernández-Lerma and J. B. Lasserre. *Further topics on discrete-time Markov control processes*, volume 42 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1999.

[Hoc16] A. Hochart. An accretive operator approach to ergodic problems for zero-sum games. In *22nd International Symposium on Mathematical Theory of Networks and Systems*, pages 315–318, Minneapolis, Minnesota, USA, July 2016. arXiv:1605.04520.

[How60] R. A. Howard. *Dynamic programming and Markov processes*. The Technology Press of M.I.T., Cambridge, Mass.; John Wiley & Sons, Inc., New York-London, 1960.

[Kar78] A. G. Kartsatos. On the equation $Tx = y$ in Banach spaces with weakly continuous duality maps. In *Nonlinear equations in abstract spaces (Proc. Internat. Sympos., Univ. Texas, Arlington, Tex., 1977)*, pages 105–112. Academic Press, New York, 1978.

[Kar81] A. G. Kartsatos. Some mapping theorems for accretive operators in Banach spaces. *J. Math. Anal. Appl.*, 82(1):169–183, 1981.

[Kat67] T. Kato. Nonlinear semigroups and evolution equations. *J. Math. Soc. Japan*, 19:508–520, 1967.

[KM97] V. N. Kolokoltsov and V. P. Maslov. *Idempotent analysis and its applications*, volume 401 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1997. Translation of ıt Idempotent analysis and its application in optimal control (Russian), "Nauka" Moscow, 1994 [MR1375021 (97d:49031)], Translated by V. E. Nazaikinskii, With an appendix by P. Del Moral.

[KN81] E. Kohlberg and A. Neyman. Asymptotic behavior of nonexpansive mappings in normed linear spaces. *Israel J. Math.*, 38(4):269–275, 1981.

[Koh74] E. Kohlberg. Repeated games with absorbing states. *Ann. Statist.*, 2:724–738, 1974.

[Koh80] E. Kohlberg. Invariant half-lines of nonexpansive piecewise-linear transformations. *Math. Oper. Res.*, 5(3):366–372, 1980.

[Kol92] V. N. Kolokoltsov. On linear, additive, and homogeneous operators in idempotent analysis. In *Idempotent analysis*, volume 13 of *Adv. Soviet Math.*, pages 87–101. Amer. Math. Soc., Providence, RI, 1992.

[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.

[KS76] J. G. Kemeny and J. L. Snell. *Finite Markov chains*. Springer-Verlag, New York-Heidelberg, 1976. Reprinting of the 1960 original, Undergraduate Texts in Mathematics.

[KS80] W. A. Kirk and R. Schöneberg. Zeros of m-accretive operators in Banach spaces. *Israel J. Math.*, 35(1-2):1–8, 1980.

[Kul97] W. Kulpa. The Poincaré-Miranda theorem. *Amer. Math. Monthly*, 104(6):545–550, 1997.

[KY92] L. A. Kontorer and S. Yu. Yakovenko. Nonlinear semigroups and infinite horizon optimization. In *Idempotent analysis*, volume 13 of *Adv. Soviet Math.*, pages 167–210. Amer. Math. Soc., Providence, RI, 1992.

[Lan71] H. Lange. *Abbildungssätze für monotone Operatoren in Hilbert- und Banach-Räumen*. PhD thesis, Freiburg (Breisgau), Univ., Math. Fak., 1971.

[Lim05] L.-H. Lim. Singular values and eigenvalues of tensors: a variational approach. In *1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pages 129–132, Puerto Vallarta, Mexico, December 2005.

[LL69] T. M. Liggett and S. A. Lippman. Stochastic games with perfect information and time average payoff. *SIAM Rev.*, 11:604–607, 1969.

[LN12] B. Lemmens and R. D. Nussbaum. *Nonlinear Perron-Frobenius theory*, volume 189 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2012.

[Mar70] R. H. Martin, Jr. A global existence theorem for autonomous differential equations in a Banach space. *Proc. Amer. Math. Soc.*, 26:307–314, 1970.

[McE06] W. M. McEneaney. *Max-plus methods for nonlinear control and estimation*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 2006.

[McE11] W. M. McEneaney. Distributed dynamic programming for discrete-time stochastic control, and idempotent algorithms. *Automatica J. IFAC*, 47(3):443–451, 2011.

[MN81] J.-F. Mertens and A. Neyman. Stochastic games. *Internat. J. Game Theory*, 10(2):53–66, 1981.

[MP15] W. M. McEneaney and A. Pandey. Development of an idempotent algorithm for a network-delay game. In *2015 Proceedings of the Conference on Control and its Applications, Paris, France, July 8-10*, pages 439–446, 2015.

[MS92] V. P. Maslov and S. N. Samborskiĭ, editors. *Idempotent analysis*, volume 13 of *Advances in Soviet Mathematics*. American Mathematical Society, Providence, RI, 1992.

[MSZ14] J.-F. Mertens, S. Sorin, and S. Zamir. *Repeated games*, volume 55. Cambridge University Press, 2014.

[MZ72] J.-F. Mertens and S. Zamir. The value of two-person zero-sum repeated games with lack of information on both sides. *Internat. J. Game Theory*, 1:39–64, 1971/72.

[Ney03] A. Neyman. Stochastic games and nonexpansive maps. In *Stochastic games and applications (Stony Brook, NY, 1999)*, volume 570 of *NATO Sci. Ser. C Math. Phys. Sci.*, pages 397–415. Kluwer Acad. Publ., Dordrecht, 2003.

[NS03] A. Neyman and S. Sorin. *Stochastic Games and Applications*, volume 570 of *Nato Science Series C*. Springer Netherlands, 2003.

[Nus88] R. D. Nussbaum. Hilbert's projective metric and iterated nonlinear maps. *Mem. Amer. Math. Soc.*, 75(391):iv+137, 1988.

[Nus89] R. D. Nussbaum. Iterated nonlinear maps and Hilbert's projective metric. II. *Mem. Amer. Math. Soc.*, 79(401):iv+118, 1989.

[Ore44] O. Ore. Galois connexions. *Trans. Amer. Math. Soc.*, 55:493–513, 1944.

[Ovc02] S. Ovchinnikov. Max-min representation of piecewise linear functions. *Beiträge Algebra Geom.*, 43(1):297–302, 2002.

[PT09] V. I. Paulsen and M. Tomforde. Vector spaces with an order unit. *Indiana Univ. Math. J.*, 58(3):1319–1359, 2009.

[PT14] A. Papadopoulos and M. Troyanov. Weak Minkowski spaces. In *Handbook of Hilbert geometry*, volume 22 of *IRMA Lect. Math. Theor. Phys.*, pages 11–32. Eur. Math. Soc., Zürich, 2014.

[Put94] M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York, 1994. A Wiley-Interscience Publication.

[Qi05] L. Qi. Eigenvalues of a real supersymmetric tensor. *J. Symbolic Comput.*, 40(6):1302–1324, 2005.

[Ren11] J. Renault. Uniform value in dynamic programming. *J. Eur. Math. Soc. (JEMS)*, 13(2):309–330, 2011.

[Ren12] J. Renault. The value of repeated games with an informed controller. *Math. Oper. Res.*, 37(1):154–179, 2012.

[Rom67] I.V. Romanovsky. Optimization of stationary control of a discrete deterministic process. *Cybernetics*, 3(2):52–62, 1967.

[RS01a] D. Rosenberg and S. Sorin. An operator approach to zero-sum repeated games. *Israel J. Math.*, 121:221–246, 2001.

[RS01b] A. M. Rubinov and I. Singer. Topical and sub-topical functions, downward sets and abstract convexity. *Optimization*, 50(5-6):307–351, 2001.

[RW98] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[Sch74] H. H. Schaefer. *Banach lattices and positive operators*. Springer-Verlag, New York-Heidelberg, 1974. Die Grundlehren der mathematischen Wissenschaften, Band 215.

[Sha53] L. S. Shapley. Stochastic games. *Proc. Nat. Acad. Sci. U. S. A.*, 39:1095–1100, 1953.

[Sio58] M. Sion. On general minimax theorems. *Pacific J. Math.*, 8:171–176, 1958.

[Sor02] S. Sorin. *A first course on zero-sum repeated games*, volume 37 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 2002.

[Sor03] S. Sorin. Symmetric incomplete information games as stochastic games. In *Stochastic games and applications (Stony Brook, NY, 1999)*, volume 570 of *NATO Sci. Ser. C Math. Phys. Sci.*, pages 323–334. Kluwer Acad. Publ., Dordrecht, 2003.

[Sor04] S. Sorin. Asymptotic properties of monotonic nonexpansive mappings. *Discrete Event Dyn. Syst.*, 14(1):109–122, 2004.

[ST13] B. Sturmfels and N. M. Tran. Combinatorial types of tropical eigenvectors. *Bull. Lond. Math. Soc.*, 45(1):27–36, 2013.

[Vig10] G. Vigeral. Evolution equations in discrete and continuous time for nonexpansive operators in Banach spaces. *ESAIM Control Optim. Calc. Var.*, 16(4):809–832, 2010.

[Vig13] G. Vigeral. A zero-zum stochastic game with compact action sets and no asymptotic value. *Dyn. Games Appl.*, 3(2):172–186, 2013.

[Whi83] P. Whittle. *Optimization over time. Vol. II.* Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 1983.

[YZ04] K. Yang and Q. Zhao. The balance problem of min-max systems is co-NP hard. *Systems Control Lett.*, 53(3-4):303–310, 2004.

[Zil16] B. Ziliotto. Zero-sum repeated games: Counterexamples to the existence of the asymptotic value and the conjecture $\mathrm{maxmin} = \lim v_n$. *Ann. Probab.*, 44(2):1107–1133, 2016.

[ZP96] U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theoret. Comput. Sci.*, 158(1-2):343–359, 1996.

Index

The page number refers to the definition of the indexed topic.

université PARIS-SACLAY

École doctorale de mathématiques Hadamard (EDMH)

Titre : Théorie de Perron-Frobenius non-linéaire et jeux stochastiques à somme nulle avec paiement moyen.

Mots clés : jeux répétés à somme nulle, jeux stochastiques à paiement moyen, contrôle ergodique, opérateurs de Shapley, applications non-expansives, opérateurs accrétifs, hypergraphes.

Résumé : Les jeux stochastiques à somme nulle possèdent une structure récursive qui s'exprime dans leur opérateur de programmation dynamique, appelé opérateur de Shapley. Ce dernier permet d'étudier le comportement asymptotique de la moyenne des paiements par unité de temps. En particulier, le paiement moyen existe et ne dépend pas de l'état initial si l'équation ergodique – une équation non-linéaire aux valeurs propres faisant intervenir l'opérateur de Shapley – admet une solution. Comprendre sous quelles conditions cette équation admet une solution est un problème central de la théorie de Perron-Frobenius non-linéaire, et constitue le principal thème d'étude de cette thèse. Diverses classes connues d'opérateur de Shapley peuvent être caractérisées par des propriétés basées entièrement sur la relation d'ordre ou la structure métrique de l'espace. Nous étendons tout d'abord cette caractérisation aux opérateurs de Shapley « sans paiements », qui proviennent de jeux sans paiements instantanés. Pour cela, nous établissons une expression sous forme minimax des fonctions homogènes de degré un et non-expansives par rapport à une norme faible de Minkowski. Nous nous intéressons ensuite au problème de savoir si l'équation ergodique a une solution pour toute perturbation additive des paiements, problème qui étend la notion d'ergodicité des chaînes de Markov. Quand les paiements sont bornés, cette propriété d'"ergodicité" est caractérisée par l'unicité, à une constante additive près, du point fixe d'un opérateur de Shapley sans paiement. Nous donnons une solution combinatoire s'exprimant au moyen d'hypergraphes à ce problème, ainsi qu'à des problèmes voisins d'existence de points fixes. Puis, nous en déduisons des résultats de complexité. En utilisant la théorie des opérateurs accrétifs, nous généralisons ensuite la condition d'hypergraphes à tous types d'opérateurs de Shapley, y compris ceux provenant de jeux dont les paiements ne sont pas bornés. Dans un troisième temps, nous considérons le problème de l'unicité, à une constant additive près, du vecteur propre. Nous montrons d'abord que l'unicité a lieu pour une perturbation générique des paiements. Puis, dans le cadre des jeux à information parfaite avec un nombre fini d'actions, nous précisons la nature géométrique de l'ensemble des perturbations où se produit l'unicité. Nous en déduisons un schéma de perturbations qui permet de résoudre les instances dégénérées pour l'itération sur les politiques.

Title : Nonlinear Perron-Frobenius theory and mean-payoff zero-sum stochastic games.

Keywords : zero-sum repeated games, mean-payoff stochastic games, ergodic control, Shapley operators, nonexpansive maps, accretive operators, hypergraphs.

Abstract : Zero-sum stochastic games have a recursive structure encompassed in their dynamic programming operator, so-called Shapley operator. The latter is a useful tool to study the asymptotic behavior of the average payoff per time unit. Particularly, the mean payoff exists and is independent of the initial state as soon as the ergodic equation – a nonlinear eigenvalue equation involving the Shapley operator – has a solution. The solvability of the latter equation is a central question in nonlinear Perron-Frobenius theory, and the main focus of the present thesis. Several known classes of Shapley operators can be characterized by properties based entirely on the order structure or the metric structure of the space. We first extend this characterization to "payment-free" Shapley operators, that is, operators arising from games without stage payments. This is derived from a general minimax formula for functions homogeneous of degree one and nonexpansive with respect to a given weak Minkowski norm. Next, we address the problem of the solvability of the ergodic equation for all additive perturbations of the payment function. This problem extends the notion of ergodicity for finite Markov chains. With bounded payment function, this "ergodicity" property is characterized by the uniqueness, up to the addition by a constant, of the fixed point of a payment-free Shapley operator. We give a combinatorial solution in terms of hypergraphs to this problem, as well as other related problems of fixed-point existence, and we infer complexity results. Then, we use the theory of accretive operators to generalize the hypergraph condition to all Shapley operators, including ones for which the payment function is not bounded. Finally, we consider the problem of uniqueness, up to the addition by a constant, of the nonlinear eigenvector. We first show that uniqueness holds for a generic additive perturbation of the payments. Then, in the framework of perfect information and finite action spaces, we provide an additional geometric description of the perturbations for which uniqueness occurs. As an application, we obtain a perturbation scheme allowing one to solve degenerate instances of stochastic games by policy iteration.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France