Algorithme de fictitious play et cycles

Richard Baron*
Université de Saint-Étienne

Jacques Durieu*
Université de Saint-Étienne

Philippe Solal*
Université de Saint-Étienne

1 Introduction

Une préoccupation récente de la théorie des jeux est de fournir une justification dynamique au concept d'équilibre de Nash. Il s'agit de modéliser les procédures mises en œuvre par les joueurs pour apprendre à jouer un équilibre de Nash dans un jeu itéré. Le point de départ est l'hypothèse que les joueurs ont une information incomplète sur la structure du jeu. Pour pallier cette incomplétude, les joueurs tirent profit de l'itération du jeu en exploitant une information factuelle rétrospective. Cette information, collectée au cours du déroulement du jeu, recense les profils de stratégies joués lors d'une suite de périodes passées retraçant une partie ou l'intégralité de l'histoire du jeu.

Suivant la manière dont les joueurs utilisent l'information factuelle rétrospective, deux classes de modèles d'apprentissage peuvent être distinguées.

1. Si l'information factuelle rétrospective est utilisée pour prédire le comportement futur des opposants, l'apprentissage est qualifié de "forward looking". À chaque période, tous les joueurs forment une nouvelle croyance sur la stratégie utilisée par chacun de leurs opposants dans l'avenir. Ensuite, compte tenu de cette croyance, chaque joueur détermine sa propre stratégie selon une règle d'adaptation plus ou moins sophistiquée.

* Les auteurs remercient Fabien Feschet pour sa suggestion concernant la preuve de la proposition 1 ainsi qu'un rapporteur anonyme pour ses commentaires.

Université de Saint-Étienne. CREUSET. 6, rue basse des rives. 42100 Saint-Étienne, France. e-mail : baron@univ-st-etienne.fr. e-mail : durieu@univ-st-etienne.fr. e-mail : solal@univ-st-etienne.fr.

2. Si l'information factuelle rétrospective est utilisée directement afin de choisir une stratégie, sans que soient formées des croyances sur les stratégies futures des opposants, l'apprentissage est qualifié de "backward looking" ; ceci parce que les joueurs sont exclusivement tournés vers le passé du jeu[1].

Notons que, contrairement aux modèles d'apprentissage de type "backward looking", les apprentissages de type "forward looking" impliquent que chaque joueur est conscient de la nature stratégique de l'environnement, i.e. que ses propres paiements dépendent des choix de stratégies des autres joueurs.

Dans cet article nous mettons l'accent sur l'apprentissage de type "forward looking"[2]. Particulièrement, nous nous focalisons sur un processus d'apprentissage qui, depuis une date récente, reçoit une attention particulière : le processus de Fictitious Play[3]. Le modèle de Fictitious Play n'est pas nouveau. Proposé par Brown (1951) et étudié par Robinson (1951), Miyasawa (1961), Shapley (1964), ce processus se présentait comme un algorithme permettant de calculer les équilibres de Nash de certains jeux. Brown considère des joueurs "fictifs" qui se comportent comme des "statisticiens". Au cours d'une itération "fictive" du jeu, chacun calcule la distribution de fréquences empiriques des stratégies jouées par ses opposants et choisit une stratégie pure optimale face à cette distribution. Mais, depuis une dizaine d'années, cet algorithme est réinterprété pour servir de justification dynamique aux équilibres de Nash. Dans ce cadre, le processus de Fictitious Play présente deux caractéristiques.

1. Tout joueur suppose que ses opposants utilisent des stratégies (mixtes) stationnaires qu'il cherche à exhiber en utilisant une information factuelle rétrospective. Plus précisément, lors de chacune des périodes d'un jeu itéré indéfiniment, les joueurs construisent, indépendamment les uns des autres, des croyances sur les stratégies mixtes utilisées par leurs opposants en calculant les fréquences empiriques des différentes stratégies pures. Puis, ils choisissent une meilleure réponse pure face à ces croyances. Autrement dit, à chaque période, les joueurs révisent leurs croyances concernant les stratégies de leurs opposants et adaptent leur stratégie.
2. L'analyse du processus de Fictitious Play peut être menée sur deux plans. Elle peut porter soit sur la suite des profils de stratégies pures joués par les joueurs soit sur la suite des croyances de chaque joueur sur les stratégies des autres joueurs.

Du point 2. émergent deux critères permettant de caractériser la convergence du processus d'apprentissage de Fictitious Play. Tout d'abord, le critère de convergence du processus de Fictitious Play peut être lié à l'étude de la suite de croyances des joueurs. Dans ce cas, le processus de Fictitious

---

[1] Cette catégorie renvoie essentiellement aux dynamiques d'imitation et de renforcement opérant.

[2] Pour une revue critique de cette littérature, cf. notamment Walliser (1998) et Sobel (2000).

[3] De manière non exhaustive, on peut mentionner les travaux de Milgrom et Roberts (1991), Fudenberg et Kreps (1993), Monderer et Shapley (1996), Foster et Young (1998).

Play converge si la suite des croyances des joueurs converge vers un profil de Nash (en stratégies pures ou mixtes). On dit que le processus de Fictitious Play *converge en croyances.* Toutefois, ce critère de convergence se révèle plutôt mal adapté lorsque l'on interprète le processus de Fictitious Play en termes d'apprentissage d'un équilibre de Nash. Ce point a été soulevé notamment par Fudenberg et Kreps (1993), Jordan (1993) et Young (1993). Des exemples de jeux de coordination 2×2 montrent que le processus de Fictitious Play peut converger en croyances (vers l'équilibre de Nash en stratégies mixtes) sans qu'un équilibre soit joué (voir Figure 1 plus bas). Autrement dit, le processus peut converger en croyances alors que les stratégies des joueurs sont perpétuellement non coordonnées. Aussi, lorsque l'objectif du modélisateur est de montrer que les joueurs peuvent apprendre à jouer un équilibre de Nash, ce critère de convergence en croyances se révèle insuffisant.

Cette remarque motive le passage à un autre critère de convergence. À la suite de Fudenberg et Kreps (1993) et Monderer et Sela (1997), le critère est la convergence de la suite des stratégies pures jouées par les joueurs. Dans ce cas, le processus de Fictitious Play converge si, à partir d'une certaine période, le profil des stratégies pures joué forme un équilibre de Nash. On dit que le processus de Fictitious Play *converge en stratégies*[4].

Contrairement au critère de convergence en croyances qui est *nécessaire mais non suffisant* pour garantir qu'un équilibre de Nash soit effectivement joué, le critère de convergence en stratégies est *nécessaire et suffisant* pour assurer que les estimations des joueurs et les profils de stratégies pures convergent vers un équilibre de Nash.

Certains résultats de convergence en stratégies ont été récemment mis en évidence. Particulièrement, Sela et Herreiner (1999) montrent que, dans un jeu de coordination 2×2, si les paiements sont des rationnels, alors le processus de Fictitious Play converge en stratégies après un nombre fini de périodes. Toutefois, cette affirmation repose sur une définition spécifique du processus de Fictitious Play qui exclut les croyances initiales des joueurs. Or, l'intégration de ces croyances est de nature à modifier ce résultat. Pour s'en persuader, il suffit de prendre l'exemple suivant.
Soit un jeu de coordination à deux joueurs. Chaque joueur dispose de deux stratégies pures : -1 et $+1$. La matrice de paiement est la suivante :

	-1	$+1$
-1	$(1,1)$	$(0,0)$
$+1$	$(0,0)$	$(1,1)$

On suppose que le joueur 1 (en ligne) croit initialement, de manière arbitraire, que le joueur 2 (en colonne) jouera la stratégie -1 avec une probabilité $\frac{1}{3}$ et la stratégie $+1$ avec la probabilité complémentaire $\frac{2}{3}$. Le joueur 2

[4] Monderer et Sela (1997) introduisent la notion de Propriété de Fictitious Play en Stratégies (PFPS). Un jeu possède la PFPS si tout processus de Fictitious Play converge vers un équilibre de Nash en stratégies pures.

a des croyances initiales symétriques. Sous ces conditions, à la période $t = 1$, le joueur 1 choisit la stratégie $+1$ et le joueur 2 la stratégie -1. Les résultats de Monderer et Shapley (1996) montrent que, dans ce jeu, la suite de croyances converge vers l'un des équilibres de Nash. La figure 1 indique que, sur deux cents itérations, la suite des croyances d'un joueur converge vers l'équilibre de Nash en stratégies mixtes du jeu, mais sans jamais pouvoir l'atteindre. Pour voir ceci, procédons par contradiction en supposant que les croyances d'un joueur atteignent, à une période quelconque, l'équilibre de Nash en stratégies mixtes. Pour les croyances du joueur 1 (respectivement du joueur 2) sur la stratégie -1 (respectivement la stratégie $+1$), on peut écrire :

$$\frac{\frac{1}{3} + a}{b} = \frac{1}{2}$$

avec $a, b \in I\!N$ où a représente le nombre de fois où la stratégie -1 (respectivement la stratégie $+1$) est observée et b est le nombre de périodes jouées. Après calculs, on obtient :

$$a = \frac{3b - 2}{6}$$

Or, ceci est en contradiction avec $a \in I\!N$. En effet, $3b$ est un multiple de 3. Donc, $3b - 2$ n'est pas un multiple de 3 ni, par conséquent, un multiple de 6. Ainsi, les joueurs n'ont jamais l'opportunité d'utiliser la règle de partage entre plusieurs meilleures réponses. Au total, les joueurs sont perpétuellement en situation de non coordination : le processus de Fictitious Play ne converge pas en stratégies bien que la suite de croyances converge vers l'équilibre de Nash en stratégies mixtes.

Cet exemple souligne la nécessité d'une analyse de la convergence en stratégies construite en liaison avec les croyances initiales adoptées par les joueurs. Particulièrement, dans la suite de cet article, nous distinguerons deux classes de croyances initiales : soit les croyances initiales des deux joueurs forment un profil de stratégies pures soit elles forment un profil de stratégies mixtes. Notre propos est de montrer que la convergence en stratégies du processus de Fictitious Play dans les jeux de coordination 2×2 dépend de manière cruciale de la classe de croyances initiales retenue. Précisément, nous établissons que la convergence en stratégies est assurée, après un nombre fini de périodes, dans toutes les catégories de jeux de coordination lorsque les croyances initiales des joueurs forment un profil de stratégies pures. En revanche, lorsque les croyances initiales des joueurs forment un profil de stratégies mixtes, les résultats en termes de convergence en stratégies sont limités à certaines catégories de jeux.

Le papier est construit de la manière suivante. Dans la section 2, nous présentons le cadre formel du modèle. Dans la section 3, nous établissons la convergence en stratégies lorsque les croyances initiales des joueurs forment un profil de stratégies pures. Dans la section 4, nous repérons les catégories de jeux de coordination pour lesquelles la convergence en stratégies est as-

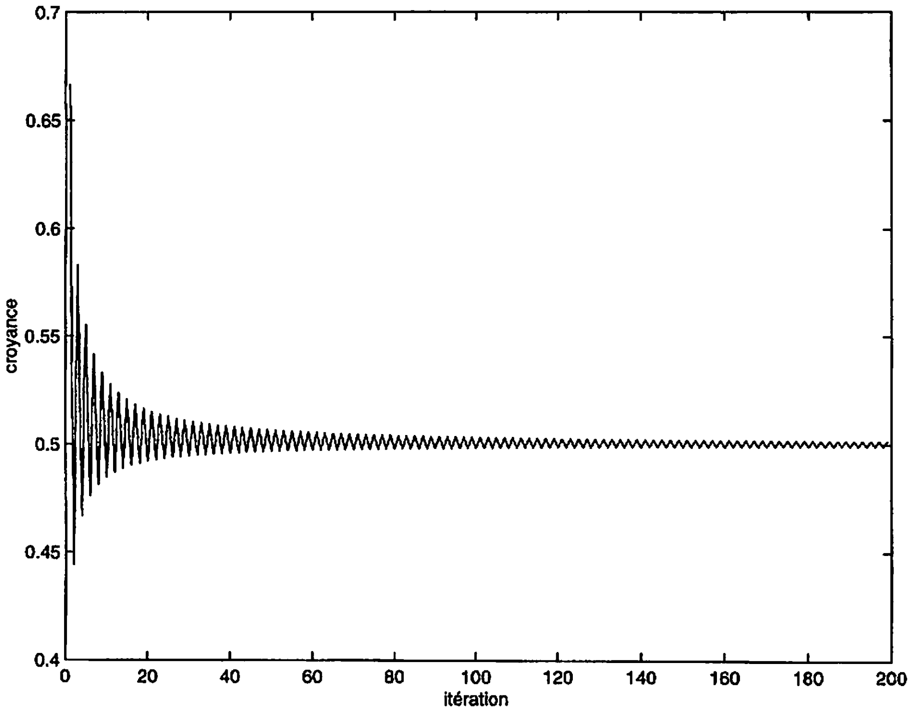

Figure 1

surée lorsque les croyances initiales forment un profil de stratégies mixtes. La section 5 conclut l'article.

2 Processus de Fictitious Play à deux joueurs

Soit Γ l'ensemble des jeux de coordination 2×2 sous forme stratégique. Dans un jeu de coordination, les joueurs ont le même nombre de stratégies pures qui peuvent être indicées de telle sorte qu'un équilibre de Nash strict soit réalisé lorsque les deux joueurs choisissent une stratégie pure ayant le même indice.

Soit $\{I, (S_i, \pi_i)_{i \in I}\}$ un jeu appartenant à Γ avec $I = \{1, 2\}$ l'ensemble des joueurs; $S_i = \{\pm 1\}$ l'ensemble des stratégies pures du joueur i ($i = 1, 2$). Un profil de stratégies pures est noté $s \in S_1 \times S_2$. Chaque joueur i est muni d'une fonction de paiement $\pi_i : S_1 \times S_2 \to Q$ où Q désigne l'ensemble des rationnels[5]. Nous considérons des jeux dans lesquels les joueurs n'ont pas

[5] Cette restriction imposée sur l'ensemble des paiements permet de traiter des cas où la convergence en stratégies repose sur l'occurence de situations d'indifférence des joueurs entre plusieurs meilleures réponses (cf. ci-dessous la preuve de la proposition 1 pour les jeux de la catégorie 1).

nécessairement la même fonction de paiement.
Soit $\Delta(S_i)$ le simplexe des stratégies mixtes du joueur i. Autrement dit, on a $\Delta(S_i) = \{x_i = (x_i(-1), x_i(+1)) \; : \; x_i(k) \geqslant 0 \, \forall k \in S_i \text{ et } \sum_{k \in S_i} x_i(k) = 1\}$. La stratégie pure $s_i = k, k \in S_i$, est le sommet k de $\Delta(S_i)$, i.e. la stratégie mixte qui attribue la probabilité 1 à $s_i = k$. On désigne par s^* un équilibre de Nash strict du jeu considéré et $x^* = (x_1^*, x_2^*)$ l'équilibre de Nash mixte.

Le processus de Fictitious Play décrit une situation dans laquelle le joueur i ($i = 1,2$) agit comme s'il faisait face à une stratégie mixte stationnaire, mais inconnue, jouée par le joueur $-i$. À une période quelconque $t \in I\!N$, le joueur i observe l'historique des choix et révise sa croyance sur la stratégie mixte censée être jouée par $-i$ en calculant la distribution empirique des stratégies pures de $-i$. Ensuite, i choisit une meilleure réponse pure face à la nouvelle estimation de la stratégie mixte de $-i$.

La définition d'un processus de Fictitious Play intègre des pondérations initiales accordées à chaque stratégie pure de l'opposant, appelées croyances initiales, et une règle de partage utilisée pour sélectionner une meilleure réponse parmi les deux. Par la suite, on suppose que si un joueur i est indifférent entre les deux meilleures réponses à la période t, il retient chacune d'entre elles avec une probabilité strictement positive. La croyance initiale d'un joueur i sur le choix s_{-i} de son opposant $-i$ à la période 0 est une fonction de pondération $b_i^0 : S_{-i} \to [0,1]$ telle que $b_i^0(-1) + b_i^0(+1) = 1$. Un historique des choix d'une longueur $t > 1$ est noté $\zeta_1^t = \{s^\iota\}_{\iota=1}^t$, avec $\zeta_1^t \in (S_1 \times S_2)^t$. À la période t, chaque joueur i révise ses croyances en ajoutant 1 au poids correspondant à la stratégie pure s_{-i} jouée en t par son opposant. À la fin de la période t, le joueur i croit que $-i$ utilise une stratégie mixte $b_i(\zeta_1^t) = (b_i(\zeta_1^t, -1), b_i(\zeta_1^t, +1))$ où $b_i(\zeta_1^t) \in \Delta(S_{-i})$,

$$b_i(\zeta_1^t, s_{-i}) = \frac{b_i^0(s_{-i}) + \nu_i(\zeta_1^t, s_{-i})}{t+1},$$

avec $\nu_i(\zeta_1^t, s_{-i}) = \sum_{\iota=1}^t x_{-i}^\iota(s_{-i})$, x_{-i}^ι étant un sommet de $\Delta(S_{-i})$ et $s_{-i} = -1, +1$. Au début de la période suivante, chaque joueur choisit une stratégie qui maximise son paiement espéré. La règle de meilleure réponse pour le joueur i est une correspondance $\mathcal{BR}_i \; : \; \Delta(S_{-i}) \twoheadrightarrow \Delta(S_i)$ qui associe à chaque croyance $b_i(\zeta_1^t) \in \Delta(S_{-i})$ l'ensemble fini non vide $\mathcal{BR}_i(b_i(\zeta_1^t))$.

3 Croyances initiales sous forme de profil de stratégies pures

Définition 1 *Une suite de profils purs est une suite $\{s^t\}_{t=1}^\infty$. On dit que le processus de Fictitious Play converge en stratégies s'il existe $T \in I\!N$ tel que, pour tout $t \geqslant T$, $s^t = s^*$.*

Dans les propositions suivantes, nous utiliserons le principe de stabilité énoncé par Fudenberg et Kreps (1993) et Monderer et Sela (1997), l'algorithme de division et le théorème de Bezout. Par la suite, Z désigne l'ensemble des entiers relatifs.

Principe de stabilité *Si un équilibre de Nash strict est joué à une période quelconque t du jeu alors il est joué lors de toutes les périodes ultérieures.*

Théorème (algorithme de division) *Pour deux entiers a et b avec $b > 0$, il existe deux uniques entiers q et r tels que $a = bq + r$ avec $0 \leqslant r < b$.*

Soient deux entiers a et b. Alors a divise b (ce que l'on écrit $a|b$) s'il existe un entier c tel que $ac = b$. À présent, rappelons que deux entiers a et b sont premiers entre eux si leur plus grand commun diviseur est 1.

Théorème 1 *Soient a et b appartenant à $Z\backslash\{0\}$ et c un diviseur du produit ab. Si a et c sont premiers entre eux alors $c|b$.*

Preuve : comme a et c sont premiers entre eux, il existe $u, v \in Z$ tels que

$$uc + va = 1 \quad \text{(Théorème de Bezout)}$$

et $ucb + vab = b$. Et, on a $c|ucb$. Puisque, par hypothèse, $c|ab$, alors $c|vab$. Donc, $c|b$. •

Proposition 1 *Pour tout jeu appartenant à Γ, le processus de Fictitious Play converge en stratégies si la croyance initiale des joueurs est une paire de stratégies pures.*

Preuve : tout d'abord, notons que si les croyances initiales forment une paire de stratégies pures affectant des pondérations strictement positives aux stratégies pures de même indice, alors un équilibre de Nash strict est joué à la période $t = 1$. D'après le principe de stabilité, on sait que le processus de Fictitious Play converge en stratégies. À présent, on considère les cas dans lesquels les croyances initiales des joueurs ne sont pas coordonnées. On décompose la preuve en trois parties, chacune correspondant à une catégorie de jeux de coordination. La catégorie 1 comprend notamment les jeux dans lesquels la matrice de paiement du joueur i contient des paiements identiques à ceux figurant dans la matrice du joueur $-i$ mais avec une permutation des coordonnées de la diagonale descendante de la matrice de $-i$. La catégorie 2 correspond notamment aux jeux dans lesquels la matrice de paiement du joueur i est la transposée de la matrice de paiement du joueur $-i$, i.e. les paiements figurant sur la diagonale montante de la matrice de i sont permutés relativement à la matrice de $-i$. La catégorie 3 contient les jeux non recensés dans les catégories 1 ou 2.[6]

Catégorie 1 : $x_i^*(+1) = x_{-i}^*(-1)$ avec $x_i^*(+1) \neq 0,5$

Nous montrons que tant que les joueurs choisissent des stratégies pures non coordonnées, ils font face de manière répétée à des situations d'indifférence entre les deux meilleures réponses. Étant donnée la règle de partage, à long terme, la probabilité de jouer un équilibre de Nash strict tend alors vers un.

6 Un exemple de jeu appartenant à chacune des trois catégories est donné en appendice.

Puisque les paiements du jeu sont des rationnels, le poids accordé à chaque stratégie pure dans l'équilibre de Nash mixte s'exprime comme un rapport d'entiers premiers entre eux. Donc, on a

$$x_i^*(+1) = x_{-i}^*(-1) = \frac{w}{t^*},$$

où w, t^* sont des entiers premiers entre eux et $w < t^*$. D'après l'algorithme de division, on peut écrire

$$\begin{aligned} \frac{w}{t^*} &= \frac{w}{qw + r} \\ &= \frac{1}{q + \frac{r}{w}} \end{aligned}$$

En multipliant le numérateur et le dénominateur par un entier strictement positif k, on obtient

$$\frac{w}{t^*} = \frac{k}{kq + \frac{kr}{w}}$$

Supposons que k soit le poids total accordé par le joueur i (resp. le joueur $-i$) à la stratégie -1 (resp. la stratégie $+1$); alors $x_i^*(-1)$ (resp. $x_{-i}^*(+1)$) est considéré comme une fonction de k. Ceci est possible puisque l'évolution des suites de croyances (et donc des suites de stratégies) de chaque joueur est complètement symétrique. Quand le joueur i choisit s_i, le joueur $-i$ choisit $-s_i$ et ainsi de suite. Pour cette raison, on peut réduire l'analyse au joueur i. Supposons que $b_i^0(-1) = k = 1$. À la période suivante, le joueur i choisit la stratégie -1. À la fin de la période $t = 1$, le poids total accordé par le joueur i à la stratégie -1 est encore de 1 puisque, par symétrie, le joueur $-i$ a choisi $+1$. Cette situation se répète jusqu'à ce que

$$\frac{\nu_i(\zeta_1^t, -1) + b_i^0(-1)}{t+1} \leqslant \frac{b_i^0(-1)}{b_i^0(-1)q + \frac{b_i^0(-1)r}{w}}$$

avec $\nu_i(\zeta_1^t, -1) = 0$, $b_i^0(-1) = 1 = k$ et pour certains $t \geqslant 1$. D'où

$$t + 1 \geqslant q + \frac{r}{w}$$

Après un nombre fini de périodes, on a

(a) $t + 1 = q + \dfrac{r}{w}$ ou bien

(b) $t + 1 > q + \dfrac{r}{w}$

Le cas (a) signifie que $b_i(\zeta_1^t, -1) = x_{-i}^*(-1)$ pour un certain t. Comme q et t sont des entiers non nuls, $\frac{r}{w}$ est un entier. Puisque $0 \leqslant r < w$ ceci tient pour

$r = 0$. Donc, on a $x^*_{-i}(-1) = x^*_i(+1) = \frac{1}{q}$ et $t^* = q$. À la période $t^* = q$, les deux joueurs sont indifférents entre les deux stratégies pures. Ils tirent chacune d'elles avec une probabilité strictement positive. Donc, à la période t^* les joueurs peuvent se coordonner. Un moment de réflexion montre que tant que les joueurs ne choisissent pas des stratégies de même indice, les opportunités de coordination se produisent à chaque période multiple de q. Donc, la probabilité de ne pas atteindre un équilibre de Nash strict à une période t tend vers zéro quand t augmente. D'après le principe de stabilité, on peut conclure que le processus de Fictitious Play converge en stratégies.

Le cas (b) signifie que $b_i(\zeta_1^t, -1) < x^*_{-i}(-1)$ pour certains t. À la période $t+1$, le joueur i choisit la stratégie $+1$ et le joueur $-i$ la stratégie -1. Donc, à la fin de la période $t+1$, $\nu_i(\zeta_1^{t+1}, -1) = \nu_i(\zeta_1^t, -1) + 1$. Le joueur i révise le poids accordé à la stratégie -1 en ajoutant 1 tant que sa croyance sur cette stratégie est strictement inférieure à $x^*_{-i}(-1)$. Après un nombre suffisant de périodes $\iota \geqslant t+1$, on a

$$\frac{\nu_i(\zeta_1^\iota, -1) + b_i^0(-1)}{\iota + 1} \geqslant \frac{k}{kq + \frac{kr}{w}}$$

avec $\nu_i(\zeta_1^\iota, -1) + b_i^0(-1) = k$. D'où

$$\iota + 1 \leqslant kq + \frac{kr}{w}$$

Supposons que $\iota + 1 = kq + \frac{kr}{w}$. Comme ι est un entier, $w|kr$. Il suit que $w|k(t^* - qw)$. Ceci implique que $w|kt^*$ (et $w|kqw$). Comme w et t^* sont premiers entre eux, par le théorème 1, $w|k$. La plus petite valeur de k telle que $w|kr$ est w. Dit autrement, la croyance accordée à la stratégie -1 par le joueur i est égale à $x^*_{-i}(-1)$ si $\nu_i(\zeta_1^\iota, -1) + b_i^0(-1) = w$. Cette égalité est vérifiée pour chaque période où k est un multiple de w. Comme dans le cas (a), on peut conclure que le processus de Fictitious Play converge en stratégies.

À présent, supposons que $\iota + 1 < kq + \frac{kr}{w}$. Le poids total accordé à la stratégie -1 par le joueur i est inchangé tant que $b_i(\zeta_1^t, -1)$ est supérieur au seuil $x^*_{-i}(-1)$. Donc $b_i(\zeta_1^t, -1)$ décroît quand t augmente. Quand la croyance $b_i(\zeta_1^t, -1)$ atteint une valeur inférieure au seuil, k augmente après chaque période jusqu'à ce qu'il atteigne la valeur w. À ce stade, on retourne à la situation précédente.

Catégorie 2 : $x^*_i(-1) = x^*_{-i}(-1)$

Supposons que $b_i^0(-1) = b_{-i}^0(+1) = 1$. À la période $t = 1$, le joueur i choisit la stratégie -1 et le joueur $-i$ la stratégie $+1$. Donc, à la fin de la période $t = 1$, $b_i(\zeta_1^1, +1) = b_{-i}(\zeta_1^1, +1) = 0,5$. On peut distinguer trois cas selon la valeur de la pondération attribuée à la stratégie -1 dans l'équilibre de Nash mixte.

Cas (a) : $0,5 < x_i^*(-1)$ $(i = 1,2)$. Les deux joueurs choisissent la stratégie $+1$. Par le principe de stabilité, on sait que la stratégie $+1$ sera jouée lors de toutes les périodes ultérieures.

Cas (b) : $0,5 > x_i^*(-1)$ $(i = 1,2)$. À la période $t = 2$, les deux joueurs choisissent la stratégie -1. Par le principe de stabilité, on sait que la stratégie -1 sera toujours jouée par la suite.

Cas (c) : $0,5 = x_i^*(-1)$ $(i = 1,2)$. Les deux joueurs sont indifférents entre leurs deux stratégies pures. Ils utilisent la règle de partage. En conséquence, il existe une probabilité strictement positive que les joueurs se coordonnent. Autrement, après calculs, on voit que la suite de profils purs joués décrit un cycle de période deux tant que les joueurs ne se coordonnent pas. À chaque période $t = \iota + 2$, $\iota \in I\!N \cup \{0\}$, $b_i(\zeta_1^t, -1) = b_{-i}(\zeta_1^t, +1) = 0,5$. Donc, la probabilité de ne pas atteindre un équilibre de Nash strict à une période t tend vers zéro quand t augmente. Par le principe de stabilité, on peut conclure que le processus de Fictitious Play converge en stratégies.

Catégorie 3 : $x_i^*(-1) \neq x_{-i}^*(-1)$ et $x_i^*(-1) \neq x_{-i}^*(+1)$

En s'inspirant de Sela et Herreiner (1999), supposons, par contradiction, qu'un équilibre de Nash strict ne soit jamais joué. Ceci implique que lorsque le joueur i choisit la stratégie -1, le joueur $-i$ choisit la stratégie $+1$ $(i = 1,2)$. En conséquence, la croyance $b_i(\zeta_1^t, -1)$ converge vers la croyance $b_{-i}(\zeta_1^t, +1)$ quand t tend vers l'infini (a). Par ailleurs, on sait qu'un jeu de Γ est non-dégénéré[7] et possède la propriété de Fictitious Play, i.e. tout processus de Fictitious Play converge en croyances vers un équilibre de Nash du jeu (Monderer et Shapley, 1996). Donc, $b_i(\zeta_1^t, -1)$ est aussi proche que souhaité de $x_{-i}^*(-1)$ et $b_{-i}(\zeta_1^t, +1)$ de $x_i^*(+1)$ quand t tend vers l'infini (b). En combinant (a) et (b), et puisque $x_i^*(+1) \neq x_{-i}^*(-1)$, on obtient une contradiction.

Ceci complète la preuve. •

Notons qu'un processus de Fictitious Play sans croyances initiales peut être formellement ramené à un processus de Fictitious Play avec des croyances initiales formant un profil de stratégies pures. Il n'est donc pas surprenant que les résultats de cette section coïncident avec ceux de Sela et Herreiner (1999). Toutefois, la décomposition de la preuve de la proposition 1 en trois temps permet de traiter de l'intégralité des catégories de jeux de Γ. En particulier, la convergence en stratégies est établie pour la catégorie 1 de jeux de coordination, exclue dans le théorème B de Sela et Herreiner (1999).

[7] Un jeu 2×2 est dit non-dégénéré s'il possède la propriété suivante : $\pi_i(-1,-1) - \pi_i(+1,-1) + \pi_i(+1,+1) - \pi_i(-1,+1) \neq 0, \forall i \in I$.

4 Croyances initiales sous forme de profils de stratégies mixtes

Considérons, à présent, les cas où les croyances initiales des joueurs forment des paires de stratégies mixtes. Comme l'atteste l'exemple présenté en introduction, la convergence en stratégies du processus de Fictitious Play n'est pas assurée pour toutes les catégories de jeux de Γ. Dans cette section, on repère les catégories de jeux pour lesquelles la convergence est assurée.

Proposition 2 *Si la croyance initiale des joueurs est une paire de stratégies complètement mixtes, le processus de Fictitious Play peut ne pas converger en stratégies pour les deux catégories de jeux de* Γ *suivantes :*

(1) $x_i^*(+1) = x_{-i}^*(-1)$ *avec* $x_i^*(+1) \neq 0,5$ $(i = 1, 2)$

(2) $x_i^*(s_i) = 0,5$, *avec* $b_i^0(s_i) \neq 0,5$ $(i = 1, 2)$

Pour les autres catégories de jeux, la convergence en stratégies est assurée.

Preuve : tout d'abord, notons que si les poids initiaux attribués aux stratégies pures de même indice sont supérieurs (ou inférieurs) au seuil critique respectif de chaque joueur, alors, à la période $t = 1$, un équilibre de Nash est joué. D'après le principe de stabilité, le processus de Fictitious Play converge en stratégies. Il reste donc à traiter les cas où les poids initiaux attribués aux stratégies pures d'indice donné sont supérieurs au seuil critique d'un joueur et inférieurs au seuil critique de l'autre joueur. On décompose la preuve en trois étapes en reprenant les catégories de jeux de Γ distinguées dans la preuve de la proposition 1.

Catégorie 1 : $x_i^*(+1) = x_{-i}^*(-1)$ avec $x_i^*(+1) \neq 0,5$

Un exemple de jeu dans l'esprit de celui proposé par Fudenberg et Kreps ((1993), p.339) suffit à établir la possibilité de non convergence en stratégies dans cette catégorie.

Catégorie 2 : $x_i^*(-1) = x_{-i}^*(-1)$

Ici, nous distinguons trois cas selon la valeur de la pondération attribuée à la stratégie -1 dans l'équilibre de Nash mixte et selon les croyances initiales.

Cas (a) : $x_i^*(-1) = x_{-i}^*(-1) \neq 0,5$. Supposons que les joueurs ne se coordonnent pas au long d'une histoire finie. Supposons qu'à la période $t = 1$ le joueur i choisisse la stratégie -1 et le joueur $-i$ la stratégie $+1$. Ceci implique $b_i^0(-1) > b_{-i}^0(-1)$. À la fin de la période $t = 1$, on a

$$b_i(\zeta_1^1, -1) = \frac{b_i^0(-1)}{2} \text{ et } b_{-i}(\zeta_1^1, -1) = \frac{b_{-i}^0(-1) + 1}{2}$$

À la période $t = 2$, si les joueurs ne se coordonnent pas, $s_i = +1$ et $s_{-i} = -1$ car $b_i(\zeta_1^1, -1) < b_{-i}(\zeta_1^1, -1)$. À la fin de la période, on a

$$b_i(\zeta_1^2, -1) = \frac{b_i^0(-1) + 1}{3} \text{ et } b_{-i}(\zeta_1^2, -1) = \frac{b_{-i}^0(-1) + 1}{3}$$

D'où, $b_i(\zeta_1^2,-1) > b_{-i}(\zeta_1^2,-1)$. À la période $t = 3$, si les joueurs ne se coordonnent pas, $s_i = -1$ et $s_{-i} = +1$ car $b_i(\zeta_1^2,-1) > b_{-i}(\zeta_1^2,-1)$. D'une manière générale, un calcul montre que, pour une période t, on a

$$b_i(\zeta_1^t,-1) = \begin{cases} \dfrac{b_i^0(-1)+\frac{t}{2}}{t+1} & \text{si } t \text{ est pair} \\[2ex] \dfrac{b_i^0(-1)+\frac{t-1}{2}}{t+1} & \text{si } t \text{ est impair} \end{cases} \tag{1}$$

et

$$b_{-i}(\zeta_1^t,-1) = \begin{cases} \dfrac{b_{-i}^0(-1)+\frac{t}{2}}{t+1} & \text{si } t \text{ est pair} \\[2ex] \dfrac{b_{-i}^0(-1)+\frac{t+1}{2}}{t+1} & \text{si } t \text{ est impair} \end{cases} \tag{2}$$

tant que les joueurs ne parviennent pas à se coordonner. Les expressions ci-dessus impliquent que $b_{-i}(\zeta_1^t,-1)$ et $b_i(\zeta_1^t,-1)$ sont proches de $0,5$ quand t augmente. Mais comme par hypothèse $x_i^*(-1) = x_{-i}^*(-1) \neq 0,5$, il doit exister une période t^* telle que soit $b_i(\zeta_1^t,-1)$ et $b_{-i}(\zeta_1^t,-1)$ sont supérieurs à $x_i^*(-1) = x_{-i}^*(-1)$ soit $b_i(\zeta_1^t,-1)$ et $b_{-i}(\zeta_1^t,-1)$ sont inférieurs à $x_i^*(-1) = x_{-i}^*(-1)$. Donc, à chaque période $\iota \geqslant t^*$ les joueurs se coordonneront. La probabilité de ne pas atteindre un équilibre de Nash tend vers zéro lorsque t augmente, ce qui est en contradiction avec l'hypothèse d'absence de coordination. Par le principe de stabilité, le processus de Fictitious Play converge en stratégies.

Cas (b) : $x_i^*(-1) = x_{-i}^*(-1) = 0,5$ et $b_i^0(-1) \neq 0,5$, $b_{-i}^0(-1) \neq 0,5$. En utilisant le résultat précédent, on peut conclure que la suite de profils purs ne converge pas vers un équilibre de Nash strict. En effet, la suite de croyances approche l'équilibre de Nash en stratégies mixtes sans jamais l'atteindre.

Cas (c) : $x_i^*(-1) = x_{-i}^*(-1) = 0,5$, $b_i^0(-1) = 0,5$ et $b_{-i}^0(-1) = 0,5$. En partant de (1) on obtient

$$b_i(\zeta_1^t,-1) = \frac{\frac{1}{2}+\frac{t}{2}}{t+1} = \frac{1}{2}$$

Donc, quand t est pair le joueur i est indifférent entre ses deux stratégies pures. Par le même raisonnement que précédemment, on peut dire que le processus de Fictitious Play converge en stratégies.

Catégorie 3 : $x_i^*(-1) \neq x_{-i}^*(-1)$ et $x_i^*(-1) \neq x_{-i}^*(+1)$

La preuve est similaire à celle de la proposition 1, catégorie 3.

Ceci complète la preuve. •

La proposition 2 révèle que le processus de Fictitious Play converge en stratégies pour un large sous-ensemble de Γ lorsque les croyances initiales forment un profil de stratégies mixtes. Seules deux catégories de jeux échappent à ce résultat. Premièrement, les jeux dont les matrices de paiement des deux joueurs sont identiques à une permutation près des coordonnées de la diagonale descendante et dont les poids accordés aux stratégies pures dans l'équilibre de Nash mixte ne sont pas tous égaux. Deuxièmement, les jeux dont les matrices de paiement des deux joueurs sont identiques à une permutation près des coordonnées de la diagonale montante et dont les poids accordés à chaque stratégie pure dans l'équilibre de Nash mixte sont égaux et différents des croyances initiales pour chaque joueur. Dans ces deux catégories de jeux de coordination, le processus de Fictitious Play peut converger en croyances vers l'équilibre de Nash mixte du jeu mais sans jamais l'atteindre. Les stratégies pures jouées par chaque joueur sont alors perpétuellement non coordonnées.

5 Conclusion

Lorsque l'on se positionne dans une perspective de justification dynamique des équilibres de Nash, le critère de convergence en stratégies est plus adapté que le critère de convergence en croyances. Toutefois, avec ce critère, la formulation de résultats se révèle assez délicate. Même dans la classe des jeux de coordination 2×2, une convergence en stratégies générale est impossible. Une analyse menée conditionnellement aux croyances initiales des joueurs permet une caractérisation complète des résultats de convergence.

Appendice

Catégorie 1 : $x_i^*(+1) = x_{-i}^*(-1)$ avec $x_i^*(+1) \neq 0,5$

	-1	$+1$
-1	$(3,4)$	$(1,1)$
$+1$	$(2,2)$	$(4,3)$

Catégorie 2 : $x_i^*(-1) = x_{-i}^*(-1)$

	-1	$+1$
-1	$(3,3)$	$(0,-1)$
$+1$	$(-1,0)$	$(4,4)$

Catégorie 3 : $x_i^*(-1) \neq x_{-i}^*(-1)$ et $x_i^*(-1) \neq x_{-i}^*(+1)$

	-1	$+1$
-1	$(3,3)$	$(1,0)$
$+1$	$(2,0)$	$(4,4)$

Reference

Brown G. (1951), "Iterative Solutions of Fictitious Play", in T.C. Koopmans (ed.), *Activity Analysis of Production and Allocation*, Wiley.

Foster D. et P. Young (1998), "On the Nonconvergence of Fictitious Play in Coordination Games", *Games and Economic Behavior*, 25, pp.79-91.

Fudenberg D. et D. Kreps (1993), "Learning Mixed Equilibria", *Games and Economic Behavior*, 5, pp.305-367.

Jordan J. (1993), "Three Problems in Learning Mixed Strategy Equilibria", *Games and Economic Behavior*, 5, pp.368-386.

Milgrom P. et J. Roberts (1991), "Adaptive and Sophisticated Learning in Normal Form Games", *Games and Economic Behavior*, 3, pp.82-100.

Miyasawa K. (1961), "On the convergence of the learning process in a 2×2 non-zero-sum two person game", Research memorandum, 33, Princeton University.

Monderer D. et A. Sela (1997), "Fictitious Play and No-cycling Conditions", Mimeo, University of Mannheim.

Monderer D. et A. Sela (1996), "A 2×2 Game without the Fictitious Play Property", *Games and Economic Behavior*, 14, pp.144-148.

Monderer D. et L. Shapley (1996), "Fictitious Play Property for Games with Identical Interests" *Journal of Economic Theory*, 68, pp.258-265.

Robinson J. (1951), "An Iterative Method of Solving a Game", *Annals of Mathematics*, 54, pp.296-301.

Sela A. et D. Herreiner (1999), "Fictitious Play in Coordination Games", *International Journal of Game Theory*, 28, pp.189-198.

Shapley L. (1964), "Some Topics in Two-person Games", in M. Dresher, L. Shapley, A. Tucker (eds.), *Advances in Game Theory*, Princeton University Press, Princeton.

Sobel J. (2000), "Economists' Models of Learning", *Journal of Economic Theory*, 94, pp.241-262.

Young P. (1993), "The Evolution of Conventions", *Econometrica*, 61, pp.57-84.

Walliser B. (1998), "A Spectrum of Equilibrium Processes in Games" *Journal of Evolutionary Economics*, 8, pp.67-87.