Le critère de stabilité stochastique en théorie des jeux évolutionnaires

Richard Baron et Philippe Solal

Édition électronique

URL : <https://journals.openedition.org/rei/381>

DOI : 10.4000/rei.381

ISSN : 1773-0198

Éditeur

De Boeck Supérieur

Édition imprimée

Date de publication : 15 septembre 2006

Pagination : 141-160

ISSN : 0154-3229

Référence électronique

Richard Baron et Philippe Solal, « Le critère de stabilité stochastique en théorie des jeux évolutionnaires », *Revue d'économie industrielle* [En ligne], 114-115 | 2e-3e trimestre 2006, mis en ligne le 03 décembre 2007, consulté le 04 juin 2022. URL : <http://journals.openedition.org/rei/381> ; DOI : <https://doi.org/10.4000/rei.381>

LE CRITÈRE DE STABILITÉ STOCHASTIQUE EN THÉORIE DES JEUX ÉVOLUTIONNAIRES

Mots-clés : Évolution, Fictitious Play, imitation, théorie des jeux, stabilité stochastique.

Key words : Evolution, Fictitious Play, Imitation, Game Theory, Stochastic Stability.

I. — INTRODUCTION

Afin de modéliser l'évolution des comportements d'une population d'agents en situation d'interaction stratégique, la théorie des jeux, dans sa version évolutionnaire, formule le cadre suivant. Une population d'agents est appelée à jouer de manière récurrente un jeu sous forme stratégique. Les agents reçoivent une information incomplète sur la structure du jeu et leur capacité à combler cette incomplétude à l'aide de l'information factuelle enregistrée au cours du temps est nulle. La récurrence du jeu ouvre simplement la possibilité aux agents d'ajuster leur comportement en fonction de croyances qu'ils peuvent former à partir de l'histoire du jeu sur les choix de leurs opposants. L'approche évolutionnaire se singularise par l'hypothèse selon laquelle les agents n'ont pas la capacité à former des croyances sophistiquées sur le comportement de leurs opposants. Ceci implique qu'un agent n'est pas en mesure d'apprendre le fonctionnement du système d'interactions dans lequel il évolue et ne recherche donc pas à reconstruire un modèle de comportements de ses opposants. Il se contente de croyances statiques ou myopes dans le sens où celles-ci reconduisent pour la période suivante les choix passés des opposants (soit les stratégies de la dernière période, soit un résumé statistique des stratégies jouées sur plusieurs périodes écoulées). Ensuite, l'agent utilise cette croyance pour prendre une décision, par exemple en choisissant une stratégie optimale étant donné cette croyance. La connaissance des fonctions de gain des adversaires n'est pas utile tant que l'agent se contente de cette règle d'optimisation myope. Dans le cas limite où l'agent ne forme aucune croyance sur l'évolution des comportements de ses adversaires, il applique mécaniquement une règle de décision simple comme « imiter la stratégie qui m'a procuré le gain le plus élevé jus-

qu'à cette période ». Ce type d'imitation ne requiert pas que l'agent soit informé des fonctions de gain de ses adversaires ou même du nombre et de l'identité des adversaires auxquels il fait face. Il suffit qu'il puisse garder en mémoire les gains obtenus dans le passé et les stratégies associées à ces gains. Lorsque les agents s'en tiennent à ce type de comportement élémentaire, il n'est pas nécessaire qu'ils aient conscience de la nature stratégique de leur environnement.

Ce qui intéresse la théorie des jeux évolutionnaires est la question de la convergence de ces processus sociaux adaptatifs : est-ce que le comportement de la population d'agents converge vers un état qui correspond à un équilibre de Nash ? Et si oui lequel ? Si le jeu ne possède pas d'équilibre de Nash en stratégies pures, quel est le comportement de la population susceptible d'émerger à long terme ? Puisque les agents ne sont pas en mesure de combler leur déficit d'information stratégique au cours du temps, la convergence éventuelle du processus vers un équilibre de Nash ne signifie pas que les agents apprennent à jouer cet équilibre, mais plutôt que l'équilibre de Nash est le produit non intentionnel d'interactions répétées intentionnelles pour utiliser une formule hayekienne. Donc, contrairement aux processus de contagion épistémique présentés dans la première partie de ce numéro spécial, les processus évolutionnaires modélisent des apprentissages sociaux où la répétition effective des interactions et les adaptations à l'évolution des comportements observés se substituent aux processus virtuels d'interaction dans lesquels les agents apprennent mentalement à simuler le comportement de leurs adversaires (sur ce point cf., Walliser, 1998).

L'objectif de cet article est de présenter le critère de stabilité stochastique et de l'appliquer à trois modèles de décision : le modèle de Fictitious Play, un modèle d'imitation et le modèle logit. Les potentialités du modèle logit, notamment dans le cadre d'interactions spatiales, sont exploitées plus avant dans l'article de Berninghaus, Haller et Outkin. Le critère de stabilité stochastique a été introduit en théorie des jeux par Foster et Young (1990). Un état d'un système soumis à un processus d'évolution adaptatif est stochastiquement stable si sa résistance à un processus de perturbations d'amplitudes variables, mais de faible occurrence, est optimale dans ce système. Ainsi, une histoire tronquée du jeu (une suite de longueur *a priori* fixée de profils de stratégies) sera stochastiquement stable relativement à un modèle d'évolution (imitation, Fictitious Play, etc.) si sa résistance face à un certain processus stochastique de perturbations est la plus développée du système. Cette résistance optimale signifie que le processus visite plus fréquemment cette histoire que les autres histoires du système de sorte que, sur le long terme, la fraction de temps passée dans cette histoire devient prépondérante, et, dans ce sens-là, on peut dire que ce moment d'histoire du jeu est le résultat de la sélection opérée par le processus d'évolution. La construction de modèles dans lesquels les histoires sélectionnées sont constituées d'un même équilibre de Nash ou, lorsque le jeu n'en possède pas, d'un ensemble de stratégies qui correspond à une généralisation ensembliste de l'équilibre de Nash a bien évidemment un inté-

rêt théorique fort. L'article est organisé de la manière suivante. La section 2 est consacrée aux définitions et à la présentation de certains concepts ensemblistes. La section 3 propose un cadre général d'un modèle d'évolution qui intègre un processus stochastique de perturbations. Trois modèles spécifiques d'évolution distincts sont ensuite présentés (modèle logit, modèle de Fictitious Play, modèle d'imitation). Les deux sections suivantes présentent une méthode (Solberg, 1975, Young, 1993) pour calculer les distributions stationnaires de ces processus et comparent les résultats de sélection obtenus à l'aide de cette méthode par chacun de ces modèles respectivement.

II. — CORRESPONDANCES COMPORTEMENTALES

On considère un jeu fini à $n \in \mathbb{N}$ joueurs sous forme stratégique,

$$G = (N, (X_i)_{i \in N}, (u_i)_{i \in N}),$$

$N = \{1, 2, \dots, n\}$ désigne l'ensemble fini des n joueurs. Pour chaque $i \in N$, X_i est l'ensemble fini des stratégies (pures) à la disposition de i . On note $X_{-i} = \prod_{j \in N \setminus \{i\}} X_j$ l'ensemble des profils de stratégies des opposants au joueur $i \in N$, et $X = \prod_{i \in N} X_i$ est l'ensemble des profils de stratégies du jeu. On écrira également (x_{-i}, x_i) au lieu de $x \in X$. Chaque joueur a des préférences sur les profils $x \in X$ représentées par une fonction de gain $u_i : X \rightarrow \mathbb{R}$ qui associe à chaque $x \in X$ un unique gain réel $u_i(x) \in \mathbb{R}$.

Pour chaque $j \in N$, $\Delta(X_j)$ désigne l'ensemble des distributions de probabilité sur X_j . Nous interprétons chaque élément $q_i = (q_i(x_{-i}))_{x_{-i} \in X_{-i}}$ de $\Theta(X_{-i}) = \prod_{j \neq i} \Delta(X_j)$ comme une croyance de i sur les choix de ses opposants. Lorsque la croyance du joueur i est concentrée sur un unique profil $x_{-i} \in X_{-i}$, *i.e.* il existe un unique $x_{-i} \in X_{-i}$ tel que $q_i(x_{-i}) = 1$, on dira que sa croyance est x_{-i} par souci de simplicité.

Nous allons maintenant introduire quelques concepts solutions ensemblistes qui généralisent le concept d'équilibre de Nash strict. Ces concepts sont utiles pour deux raisons. La première raison est que ces ensembles ont une interprétation dynamique intuitive, ce qui fait leur intérêt dans les modèles d'évolution. La seconde est que ces ensembles existent dans tout jeu fini, ce qui n'est pas le cas de l'équilibre de Nash (en stratégies pures). Ainsi, les résultats de sélection évolutionnaire (de ces ensembles) que nous exposerons dans cet article s'appliqueront à tout jeu fini. Certes, il n'est pas exclu que ces ensembles sélectionnés soient larges, et, par conséquent, que le pouvoir prédictif du modèle soit restreint. Néanmoins, nous serons toujours en mesure de prévoir que les stratégies qui n'appartiennent pas à ces ensembles auront peu de chance d'être observées à long terme. Heureusement, dans nombre de cas, un modèle d'évolution sélectionne un équilibre de Nash strict parmi plusieurs.

La correspondance *des réponses optimales* d'un joueur $i \in N$, $R_i : \Theta(X_{-i}) \rightarrow X_i$ associe à chaque $q_i \in \Theta(X_{-i})$, l'ensemble non vide des réponses optimales du joueur i ,

$$R_i(q_i) = \{z_i \in X_i : \sum_{x_{-i} \in X_{-i}} q_i(x_{-i}) (u_i(z_i, x_{-i}) - u_i(x_i, x_{-i})) \geq 0 \quad \forall x_i \in X_i\}$$

Considérons les sous-ensembles de stratégies $Z_i \subseteq X_i$, $i \in N$, le produit cartésien $Z = \prod_{i \in N} Z_i$ et un réel $p \in [0, 1]$. Le sous-ensemble $\Theta(Z_{-i}, p) \subseteq \Theta(X_{-i})$ est l'ensemble des croyances du joueur $i \in N$ telles que le poids accordé aux profils de stratégies de Z_{-i} est supérieur ou égal à p i.e. $\sum_{z_{-i} \in Z_{-i}} q_i(z_{-i}) \geq p$.

Un produit cartésien $Z = \prod_{i \in N} Z_i$ de profils de stratégies est un *ensemble p -optimal* si pour tout $i \in N$ et pour chaque croyance $q_i \in \Theta(Z_{-i}, p)$, $R_i(q_i) \subseteq Z_i$. Un tel ensemble est minimal s'il n'existe pas $Y = \prod_{i \in N} Y_i \subset Z$ qui soit p -optimal. Chaque jeu fini G possède au moins un ensemble p -optimal minimal pour chaque $p \in [0, 1]$, et si $p \geq 1/2$ cet ensemble est unique (possiblement X) (Tercieux, 2004, Durieu, Solal, Tercieux, 2005).

Observons qu'un ensemble p -optimal est également 1-optimal, c'est-à-dire un ensemble fermé sous la correspondance de réponses optimales au sens de Basu et Weibull (1991). De plus, lorsqu'un tel ensemble est un singleton, il s'agit d'un *équilibre de Nash p -dominant en risque* au sens de Morris, Rob et Shin (1995). On peut donc voir un ensemble p -optimal comme un ensemble fermé par la correspondance de réponses optimales qui satisfait la condition supplémentaire de p -dominance en risque. Ceci signifie que si chaque joueur $i \in N$ croit que ses opposants choisiront un profil dans Z_{-i} avec un poids supérieur ou égal à p , alors les réponses optimales de i sont contenues dans Z_i . Plus p est petit plus le risque à jouer une stratégie dans Z_i est faible.

Ce concept a une interprétation dynamique assez intuitive. Imaginons que deux agents soient appelés à jouer un jeu à deux joueurs de manière récurrente. À chaque période, chacun des joueurs calcule la fréquence empirique des stratégies jouées par son opposant jusqu'à cette période et croit que cette distribution empirique est une bonne approximation de ce que peut jouer son opposant à la période suivante. Dans ce cas, la distribution calculée est assimilée à une croyance sur les choix possibles de son opposant à la période suivante. Supposons qu'à un point du temps, ce calcul aboutisse à une distribution pour chaque joueur $i = 1, 2$ qui alloue un poids total de p sur les stratégies de $Z_{-i} \subseteq X_{-i}$, où $Z_1 \times Z_2$ est un ensemble p -optimal minimal. Dans ce cas, par définition de l'ensemble p -optimal, nous pouvons conclure que la réponse optimale de $i = 1, 2$ appartient à Z_i . À la période suivante, chaque joueur révisé sa croyance et le poids total affecté aux stratégies de Z_{-i} par le joueur i augmente nécessairement, ce qui implique que sa réponse optimale face à cette nouvelle croyance est encore dans Z_i . En itérant ce raisonnement, nous voyons que ce processus de réponse optimale ne quittera plus l'ensemble produit Z tant que les joueurs se conforment à cette règle de décision, appelée la règle de Fictitious play. Nous comprenons également que la croyance initiale sur le

sous-ensemble de stratégies Z_i , $i = 1, 2$, détermine la trajectoire du processus. Seuls des écarts à cette règle (expérimentations de nouvelles stratégies, mutations), peuvent, en perturbant le processus, briser cette dépendance aux croyances initiales.

La correspondance des *réponses améliorantes* d'un joueur $i \in N$, $A_i : X_i \times \Theta(X_{-i}) \rightarrow X_i$ associe à chaque couple $(x_i, q_i) \in X_i \times \Theta(X_{-i})$ l'ensemble non vide des réponses qui améliorent le gain espéré du joueur i ,

$$A_i(x_i, q_i) = \{z_i \in X_i : \sum_{x_{-i} \in X_{-i}} q_i(x_{-i}) (u_i(z_i, x_{-i}) - u_i(x_i, x_{-i})) \geq 0\}$$

Remarquons que $x_i \in A_i(x_i, q_i)$ quelle que soit la croyance q_i , et que pour chaque $q_i \in \Theta(X_{-i})$,

$$R_i(q_i) = \bigcap_{x_i \in X_i} A_i(x_i, q_i)$$

Considérons les correspondances suivantes : pour chaque $i \in N$,

$D_i : X \rightarrow X$ définie par

$$D_i(x) = \{z \in X : z = (z_i, x_{-i}), z_i \in R_i(x_{-i})\},$$

et $B_i : X \rightarrow X$ définie par

$$B_i(x) = \{z \in X : z = (z_i, x_{-i}), z_i \in A_i(x_i, x_{-i})\}$$

L'ensemble $D_i(x)$ est l'ensemble des profils de stratégies qui peuvent être atteints à partir du profil $x \in X$ lorsque le joueur i choisit une réponse optimale face à x_{-i} . L'ensemble $B_i(x)$ est l'ensemble des profils de stratégies qui peuvent être atteints à partir du profil $x \in X$ lorsque le joueur i choisit une réponse améliorante au profil $x \in X$. Enfin,

$D : X \rightarrow X$ définie par

$$D(x) = \left(\bigcup_{i \in N} D_i(x) \right) \cup \{x\},$$

et $B : X \rightarrow X$ définie par

$$B(x) = \bigcup_{i \in N} B_i(x)$$

$D(x) [B(x)]$ est l'ensemble des profils de stratégies qui peuvent être atteints à partir du profil $x \in X$ si au plus un joueur choisit une réponse optimale [une réponse améliorante] face au choix de ses opposants. Évidemment, $D(x) \subseteq B(x)$.

Un sous-ensemble $Z \subseteq X$ est *fermé par des réponses optimales asynchrones* si pour chaque $z \in Z$, $D(z) \subseteq Z$. Un tel sous-ensemble est minimal s'il n'existe pas $Y \subset Z$ qui soit fermé par des réponses optimales asynchrones (Baron *et al.* 2002).

Un sous-ensemble $Z \subseteq X$ est *fermé par des réponses améliorantes asynchrones* si pour chaque $z \in Z$, $B(z) \subseteq Z$. Un tel sous-ensemble est minimal s'il n'existe pas $Y \subset Z$ qui soit fermé par des réponses améliorantes asynchrones (Josephson, Matros, 2004). Notons que ces concepts ne nécessitent pas que Z soit le produit cartésien d'ensembles de stratégies.

Remarquons que X est un ensemble fermé par des réponses améliorantes [optimales] asynchrones. Puisque X est fini, G contient au moins un ensemble minimal fermé par des réponses améliorantes [optimales] asynchrones.

Là encore, ces deux concepts sont utiles lorsque l'on souhaite modéliser des processus d'évolutions de stratégies. Supposons que chaque agent d'une population soit désigné pour jouer un même rôle, disons le rôle i , dans un jeu sous forme stratégique. Supposons que pour chaque période, un unique agent, disons i , reçoive l'opportunité d'adapter sa stratégie et se satisfasse d'une stratégie qui améliore son gain courant face au choix de ses opposants. Précisément, à la période t où cet agent peut s'adapter à l'environnement, il observe le profil des choix x_{-i}^{t-1} de la période précédente et estime, à juste titre d'ailleurs, que ce profil n'évoluera pas à la période suivante, puis il choisit chaque stratégie $z_i^t \in B_i(x_{-i}^{t-1})$ avec une probabilité strictement positive. Ce faisant, il ne réalise pas que les autres agents réagiront à sa déviation lorsqu'ils en auront l'opportunité dans le futur. Il suppose donc que son environnement est stationnaire et se satisfait d'une stratégie qui améliore son gain courant. À l'étape suivante, le processus transite vers un état $x^t = (z_i^t, x_{-i}^{t-1})$, où $z_i^t \in B_i(x_{-i}^{t-1})$ conditionnellement au fait que i reçoive l'opportunité d'adapter sa stratégie à la période t . On comprend que ce processus asynchrone de réponses améliorantes converge vers des ensembles fermés par des réponses améliorantes minimaux. Enfin, si tous ces ensembles minimaux sont des singletons (des équilibres de Nash stricts), on peut conclure qu'à partir de n'importe quel profil initial ce processus converge vers un équilibre de Nash (strict) du jeu. En ce sens, la population apprend à jouer un équilibre de Nash strict bien que chacun de ses membres ne possède pas une information complète sur la structure du jeu. Cette information, dispersée entre les joueurs, n'est pas un réel obstacle à l'émergence d'un équilibre de Nash strict à condition que les interactions soient suffisamment nombreuses. Bien sûr, la question de la sélection entre plusieurs équilibres n'est pas tranchée. Cette question sera abordée dans les sections suivantes.

Nous terminons cette section par deux remarques. Tout d'abord, un ensemble fermé par des réponses optimales asynchrones n'est pas nécessairement un ensemble 1-optimal, et ceci, même si l'ensemble fermé par des réponses optimales asynchrones est un produit cartésien d'ensembles de stra-

tégies. Pour fixer les idées, considérons le jeu à deux joueurs tiré de Durieu, Solal, Tercieux (2005).

Exemple 1

	A	B	C	D
1	(0, 3)	(1, 1.5)	$(\sqrt{7}, 0)$	(0, 0)
2	$\left(\frac{3\sqrt{7}}{3+\sqrt{7}}, 0\right)$	(2, 2)	$\left(\frac{3\sqrt{7}}{3+\sqrt{7}}, 0\right)$	(0, 0)
3	(3, 0)	(1, 1.5)	(0, 3)	(0, 0)
4	(0, 0)	(0, 0)	(0, 0)	(1, 1)

L'unique ensemble 1/2-optimal minimal est $\{1, 2, 3\} \times \{A, B, C\}$. Le produit cartésien $\{1, 3\} \times \{A, C\}$ est un ensemble fermé par des réponses optimales asynchrones et il est minimal, mais il ne constitue pas un ensemble 1-optimal. Pour voir ceci, il suffit de noter que la stratégie 2 est une réponse optimale face à la croyance plaçant un poids de $\sqrt{7}/(3 + \sqrt{7})$ sur la stratégie A et un poids de $3/(3 + \sqrt{7})$ sur la stratégie C. L'ensemble $\{(2, B)\}$ est 1-optimal ainsi que l'ensemble $\{(4, D)\}$. On peut se demander si des gains irrationnels ont un quelconque impact sur un processus d'évolution. La réponse est positive. Reprenons la règle de Fictitious play présentée plus haut. Il est impossible que les agents construisent des croyances du type $\sqrt{7}/(3 + \sqrt{7})$ sur la stratégie A et $3/(3 + \sqrt{7})$ sur la stratégie C à partir de l'observation des choix passés de l'opposant. Ces croyances, construites à partir de fréquences d'occurrences, sont nécessairement des nombres rationnels.

Ensuite, chaque ensemble fermé par des réponses améliorantes contient un ensemble fermé par des réponses optimales. *A contrario*, un ensemble fermé par des réponses optimales n'est pas nécessairement fermé par des réponses améliorantes. En particulier, un ensemble fermé par des réponses améliorantes minimal peut contenir strictement un ensemble fermé par des réponses optimales minimal. L'exemple suivant éclaire ce point.

Exemple 2

	A	B	C
1	(-1, 1)	(3, 0)	(4, -2)
2	(1, -1)	(2, 1)	(0, -2)
3	(0, 0)	(0, 1)	(-1, -1)

L'unique ensemble fermé par des réponses optimales asynchrones minimales est $\{(1, A), (2, A), (1, B), (2, B)\}$. L'unique ensemble fermé par des réponses améliorantes asynchrones minimales est $\{(1, A), (2, A), (3, A), (1, B), (2, B), (3, B)\}$.

III. — ÉVOLUTION

Le jeu G va être joué aux périodes $t = 1, 2, \dots$; G est le *jeu récurrent*. À chaque période t , un unique joueur $i \in N$ est tiré avec la probabilité $v_i > 0$ de la population pour adapter sa stratégie à l'aide d'une règle de décision. La décision est scindée en deux étapes. Tout d'abord, le joueur i construit une *croyance* sur le choix des opposants à la période suivante en fonction de l'histoire du jeu jusqu'à cette période. Ensuite, compte tenu de cette croyance, le joueur i adapte sa stratégie courante à l'aide d'une *règle de décision perturbée*. Une règle de décision est une procédure qui calcule une stratégie mixte en fonction de la croyance du joueur sur les choix de ses opposants. Une règle de décision perturbée est une règle de décision indexée par un niveau de bruit. Ce bruit est attribué aux erreurs ou tremblements du joueur qui l'écartent de la stratégie qu'il aurait dû exécuter en se conformant à sa règle de décision. Il existe plusieurs manières de former des croyances et d'introduire le bruit. Afin d'obtenir le cadre le plus général possible, les définitions suivantes sont nécessaires. Une *histoire tronquée* ou *état* à la période $t \geq m$ est une suite $h' = (x^{t-m+1}, \dots, x^t)$ où m est la taille de la mémoire des joueurs; $H = X^m$ est l'ensemble fini des histoires de longueur m et h désignera un élément de H . Le joueur i qui a l'opportunité d'ajuster sa stratégie construit une croyance $q_i^h \in \Theta(X_{-i})$ sur le choix de ses opposants pour la période suivante en fonction de l'histoire tronquée $h \in H$. Pour $i \in N$, $\Delta(X_i)$ désigne l'ensemble des stratégies mixtes du joueur $i \in N$. Une règle de décision perturbée pour le joueur $i \in N$ est une fonction $p_i : \Theta(X_{-i}) \times \mathbb{R}_{++} \rightarrow \Delta(X_i)$ qui associe à chaque croyance $q_i^h =: q_i$ construite à partir de l'histoire $h \in H$ et chaque $\varepsilon \in \mathbb{R}_{++}$ une unique stratégie mixte $p_i^\varepsilon(q_i) = (p_i^{\varepsilon, x_j}(q_i))_{x_j \in X_j} \in \Delta(X_i)$. Nous supposons que pour tout $i \in N$,

- (i) p_i^ε est indépendante de t .
- (ii) Pour chaque $\varepsilon \in \mathbb{R}_{++}$, $p_i^\varepsilon(q_i)$ est une stratégie complètement mixte pour tout $q_i \in \Theta(X_{-i})$.
- (iii) $p_i^\varepsilon(q_i)$ est continue en $\varepsilon \in \mathbb{R}_{++}$ et $p_i^* = \lim_{\varepsilon \rightarrow 0} p_i^\varepsilon$ existe.

Exemple 3 - La règle logit

Fixons $m = 1$ de sorte que $h' = x'$ pour tout $t = 1, 2, \dots$. Le joueur $i \in N$ qui obtient l'opportunité d'adapter sa stratégie face au profil courant $x = (x_i, x_{-i}) \in X$, construit la croyance q_i^x telle que $q_i^x(x_{-i}) = 1$ et choisira chaque stratégie $z_i \in X_i$ avec la probabilité

$$p_i^{\varepsilon, z_i}(x_{-i}) = \frac{\exp[u_i(z_i, x_{-i})/\varepsilon]}{\sum_{k_i \in X_i} \exp[u_i(k_i, x_{-i})/\varepsilon]} \quad (1)$$

La principale caractéristique de cette règle est que la probabilité d'un tremblement est fonction de la perte de gains suite au choix d'une stratégie non

optimale face à x_{-i} . Autrement dit, les tremblements dépendent de l'état x . À l'aide de l'expression (1), on peut exprimer le rapport des probabilités de choix entre deux stratégies, disons z_i et l_i , de la façon suivante :

$$\frac{p_i^{\varepsilon, z_i}(x_{-i})}{p_i^{\varepsilon, l_i}(x_{-i})} = \exp \left[\frac{u_i(z_i, x_{-i}) - u_i(l_i, x_{-i})}{\varepsilon} \right]$$

D'une part, nous remarquons que le rapport des probabilités de choix dépend exponentiellement de la différence des gains. Ainsi, lorsque z_i procure un gain strictement plus élevé que l_i face à x_{-i} , c'est-à-dire $u_i(z_i, x_{-i}) > u_i(l_i, x_{-i})$, alors la probabilité de choisir z_i est plus élevée que la probabilité de choisir l_i au point x_{-i} . D'autre part, lorsque ε tend vers zéro, on conclut facilement que $p_i^{\varepsilon, l_i}(x_{-i})$ devient négligeable. Comme ce raisonnement est valable pour toute paire de stratégies appartenant à X_i , la probabilité de choisir une stratégie non optimale face à x_{-i} est négligeable à la limite. Donc, cette règle tend vers la règle de réponse optimale myope lorsque le paramètre de bruit devient négligeable. La règle de décision logit est utilisée par Blume (1993) et Baron *et alii* (2002).

Exemple 4 - Règle de Fictitious Play

Fixons $m \geq 4$, $s \leq m$, $s \in \mathbb{N}$, et $\varepsilon \in (0, 1)$. Le joueur $i \in N$ sélectionné va construire une croyance sur le choix futur de ses opposants à partir de l'histoire $h \in H$, puis, en fonction de cette croyance, jouer une réponse optimale avec une certaine probabilité qui tend vers 1 lorsque $\varepsilon \rightarrow 0$. Premièrement, le joueur i qui obtient l'opportunité d'adapter sa stratégie échantillonne s profils parmi les m disponibles dans l'histoire $h \in H$, chaque sous-ensemble de s profils ayant la même probabilité d'être tiré. Deuxièmement, le joueur i calcule la fréquence $q_{ij}^h(x_j)$ avec laquelle chaque $j \neq i$ a joué la stratégie $x_j \in X_j$ dans cet échantillon. On a $\sum_{x_j \in X_j} q_{ij}^h(x_j) = 1$. La croyance q_i^h du joueur $i \in N$ est définie par

$$q_i^h(x_{-i}) := q_i(x_{-i}) = \prod_{j \neq i} q_{ij}^h(x_j) \text{ pour tout } x_{-i} = (x_j)_{j \neq i} \in X_{-i}$$

Troisièmement, le joueur i calcule le gain espéré face à chaque couple $(., q_i) \in X_i \times \Theta(X_{-i})$. Enfin, il joue une réponse optimale $x_i \in R_i(q_i)$ avec la probabilité

$$p_i^{\varepsilon, x_i}(q_i) = \frac{1 - \varepsilon}{\#R_i(q_i)} + \frac{\varepsilon}{\#X_i} \quad (2)$$

et la probabilité de choisir une réponse non optimale est donnée par

$$p_i^{\varepsilon, x_i}(q_i) = \frac{\varepsilon}{\#X_i}, \text{ pour tout } x_i \notin R_i(q_i) \quad (3)$$

Le paramètre ε représente la probabilité avec laquelle le joueur sélectionné échoue dans l'exécution de la stratégie optimale calculée à partir de sa croyance q_i sur le choix de ses opposants à la période suivante. Cette croyance q_i est elle-même le résultat d'une procédure d'échantillonnage dans l'histoire h puis de traitement d'information de cet échantillon. Notons que la probabilité que le joueur joue une stratégie dans $R_i(q_i)$ tend vers 1 lorsque $\varepsilon \rightarrow 0$. Cependant, et contrairement à la règle logit, la probabilité de s'écarter d'une réponse optimale appartenant à $R_i(q_i)$, $\varepsilon/\#X_i$, ne dépend pas de l'état du système. Cette règle perturbée de Fictitious Play avec mémoire limitée et échantillonnage a été introduite par Young (1993).

Exemple 5 - Règle d'imitation

Fixons $m \geq 4$, $s \leq m$, $s \in \mathbb{N}$, et $\varepsilon \in (0, 1)$. Le joueur $i \in N$ qui reçoit l'opportunité d'adapter sa stratégie observe l'histoire $h \in H$, échantillonne s profils dans h de la même façon que dans l'exemple précédent. Le joueur i observe chacune des stratégies qu'il a jouées ainsi que les gains obtenus avec chacune de ses stratégies dans cet échantillon d'histoire h i.e. il observe $(x_i^1, u_i(x^1))$, $\dots$, $(x_i^s, u_i(x^s))$. Ensuite, il recherche les stratégies avec lesquelles il a reçu le gain le plus élevé dans cet échantillon :

$$S_i = \{x_i^k, k \in \{1, 2, \dots, s\} \mid u_i(x_i^k, x_{-i}^k) = \max\{u_i(x^1), \dots, u_i(x^s)\}\}$$

La règle de décision du joueur i est définie de la façon suivante : pour toute croyance $q_i^h =: q_i \in \Theta(X_{-i})$,

$$p_i^{\varepsilon, x_i}(q_i) = \frac{(1 - \varepsilon)}{\#S_i} + \frac{\varepsilon}{\#X_i} \text{ pour tout } x_i \in S_i \quad (4)$$

et

$$p_i^{\varepsilon, x_i}(q_i) = \frac{\varepsilon}{\#X_i} \text{ pour tout } x_i \notin S_i \quad (5)$$

Le paramètre ε représente la probabilité avec laquelle le joueur sélectionné échoue dans l'imitation de l'une des stratégies les plus profitables dans l'échantillon. Notons que la probabilité que le joueur i imite une stratégie dans S_i tend vers 1 lorsque $\varepsilon \rightarrow 0$. Cependant, et contrairement à la règle logit, la probabilité de ne pas imiter l'une des stratégies les plus profitables dans l'échantillon, $\varepsilon/\#X_i$, est indépendante de l'état dans lequel se trouve le processus. Enfin, le choix d'une stratégie est indépendant de la croyance du joueur sur le choix de ses opposants. En effet, le joueur se contente d'observer les stratégies qu'il a jouées ainsi que les gains obtenus avec chacune de ses stratégies dans cet échantillon d'histoire, puis tire mécaniquement une des stratégies de S_i sans plus de justification. Il y a deux interprétations possibles : soit le joueur n'utilise pas ses croyances pour aboutir à une décision par souci d'économiser des

ressources devant être mobilisées pour mettre en cohérence ses croyances et son choix d'une stratégie, soit il ne forme en réalité aucune croyance pertinente pour jouer le jeu. Cette règle d'imitation perturbée avec mémoire limitée et échantillonnage a été introduite par Josephson et Matros (2004).

En gardant en mémoire ces exemples, nous pouvons aller un peu plus loin dans la description de notre modèle d'évolution. Pour une histoire $h^t = (x^{t-m-1}, \dots, x^t) \in H$ à la période t , le processus stochastique transite vers un état de la forme $h^{t+1} = (x^{t-m-2}, \dots, x^{t+1})$ à la période suivante où $x^{t+1} = (x_{i+1}^t, x_{-i}^t)$, et $i \in N$ est le joueur qui a adapté sa stratégie. Un tel état h^{t+1} est un *successeur* de h^t . Nos hypothèses impliquent que le processus transite d'un état vers un successeur à chaque période selon $(v_i)_{i \in N}$, $(p_i^\varepsilon)_{i \in N}$ et la taille de l'échantillon s lorsque la règle de décision fait intervenir la procédure d'échantillonnage.

Ces opérations définissent une chaîne de Markov sur H ; M^ε désignera la matrice de transition de format $\#H \times \#H$. Pour chaque $\varepsilon > 0$, la chaîne de Markov M^ε est irréductible et apériodique, et admet donc une unique distribution stationnaire $(\mu_h)_{h \in H}$ i.e. un vecteur (ligne) de probabilité sur H tel que $\mu^\varepsilon = \mu^\varepsilon M^\varepsilon$. Un résultat standard établit que si M^ε est irréductible et apériodique, alors la probabilité que le processus soit dans l'état $h^t := h$, t grand, est approximativement donnée par μ_h^ε , et ceci, quel que soit l'état initial $h^1 \in H$. La principale question qui intéresse la théorie des jeux est la caractérisation du support de la distribution $\mu^* = \lim_{\varepsilon \rightarrow 0} \mu^\varepsilon$:

$$C(\mu^*) = \{h \in H : \mu_h^* > 0\}$$

Les états qui appartiennent à $C(\mu^*)$ sont appelés *états stochastiquement stables*; ils correspondent aux histoires tronquées qui sont visitées par le processus de manière récurrente lorsque le niveau de bruit dans le système est faible. En toute généralité, μ^* n'existe pas nécessairement sous les conditions (i)-(ii)-(iii). Cependant, pour les exemples 3-5 la distribution μ^* existe, et nous supposons par la suite que μ^* existe. Par (iii) $M^* = \lim_{\varepsilon \rightarrow 0} M^\varepsilon$ existe, et

$$\lim_{\varepsilon \rightarrow 0} \mu^\varepsilon = \lim_{\varepsilon \rightarrow 0} \mu^\varepsilon M^\varepsilon \Rightarrow \mu^* = \mu^* M^* \quad (6)$$

La matrice M^* est la matrice de transition du *processus non perturbé*. Une conséquence directe de l'implication ci-dessus est que chaque état stochastiquement appartient à un ensemble récurrent du processus non perturbé et que si un état est stochastiquement stable, alors chaque état de l'ensemble récurrent auquel il appartient est stochastiquement stable. Pour clarifier les idées, dans l'exemple 3, les ensembles récurrents du processus non perturbé sont les ensembles fermés par des réponses optimales asynchrones minimaux. En effet, lorsque $\varepsilon \rightarrow 0$, la probabilité que le joueur qui adapte sa stratégie choisisse une réponse optimale face au profil courant tend vers 1, et à chaque période un seul joueur adapte sa stratégie. Le processus non perturbé associé au modèle d'évolution asynchrone logit est donc un processus adaptatif asynchrone de réponse optimale. Nous avons donc un premier résultat.

Proposition 1 - (Baron et alii, 2002)

Supposons que les joueurs adoptent la règle logit de l'exemple 3. Alors, l'ensemble des états stochastiquement stables est une réunion d'ensembles fermés par des réponses optimales asynchrones minimaux. Donc, sur chaque état stochastiquement stable au moins un joueur exécute une stratégie optimale face au choix de ses adversaires.

Exemple 6

Considérons le jeu de l'exemple 2. Par une application de la proposition 1 l'ensemble des états stochastiquement stables est $\{(1,A), (2,A), (1,B), (2,B)\}$. Pour certains jeux, le critère de sélection stochastique ne permet pas de sélectionner un profil en particulier. Considérons le jeu à deux joueurs tel que $X_1 = X_2 = \{0, 1\}$, $u_1(x_1, x_2) = 1$ et $u_2(x_1, x_2) = -1$ si $x_1 = x_2$, et $u_2(x_1, x_2) = 1$, $u_1(x_1, x_2) = -1$ si $x_1 \neq x_2$. L'unique ensemble fermé par des réponses optimales asynchrones est $X_1 \times X_2$. Par une application de la proposition 1, chaque état est stochastiquement stable. Dans ce cas simple, on peut calculer μ^* : $\mu_{(0,0)}^* = \mu_{(1,1)}^* = v_1/2$, $\mu_{(0,1)}^* = \mu_{(1,0)}^* = v_2/2$.

Pour des résultats complémentaires sur ce modèle d'évolution engendré par la règle logit, nous renvoyons le lecteur à l'article de Berninghaus, Haller Outkin dans ce numéro et à Durieu, Haller, Solal (2005). Dans ce qui suit nous allons étudier le modèle de Fictitious Play (exemple 4) et le modèle d'imitation (exemple 5).

IV. — CARACTÉRISATION DE $C(\mu^*)$

Dans les modèles de Fictitious Play et d'imitation, la perturbation est modélisée par une expérience de Bernoulli. Il est utile de ramener le problème de la recherche des éléments de $C(\mu^*)$ à un programme d'optimisation que nous allons présenter. Pour $h, h' \in H$, désignons par $M_{hh'}^\varepsilon$ la probabilité de transiter de h vers h' en une période. Le *graphe stochastique* de la chaîne de Markov est le graphe dont les sommets sont les éléments de H et les arcs sont les transitions réalisables en une période par le processus : un arc $h \rightarrow h'$ signifie que $M_{hh'}^\varepsilon > 0$ et $M_{hh'}^\varepsilon$ est le poids de l'arc $h \rightarrow h'$. Les *arbres de recouvrement* du graphe stochastique sont les arbres de racine $h \in H$ et dont les autres sommets sont les états $h' \in H \setminus \{h\}$; chaque état h' possédant un unique successeur sur l'arbre. Le symbole T_h désignera l'ensemble des arbres de recouvrement de racine $h \in H$.

Théorème 1 - (Solberg, 1975)

Fixons $\varepsilon > 0$. Pour chaque $h \in H$,

$$\mu_h^\varepsilon = \frac{\sum_{\tau \in T_h} \left(\prod_{h' \rightarrow h'' \in \tau} M_{h'h''}^\varepsilon \right)}{Z}$$

où Z est la constante de normalisation.

Ce théorème dit que chaque composante de la distribution stationnaire de M^ε est proportionnelle à la somme sur tous les arbres de recouvrement du graphe de M^ε du produit des probabilités de transition sur l'arbre. Enfin, on définit la résistance d'un arbre de recouvrement de racine h , $\tau \in T_h$, par

$$r(\tau) = \sum_{h' \rightarrow h'' \in \tau} r(h', h'')$$

où $r(h', h'') \in \{0, 1\}$ est le nombre minimal de perturbations ε pour transiter de l'état h' vers un successeur h'' .

On appelle *potentiel stochastique* v_h d'un état $h \in H$, l'arbre de recouvrement de racine $h \in H$ qui possède la résistance minimale :

$$v_h = \min_{\tau \in T_h} r(\tau)$$

Le potentiel stochastique v_h calcule le nombre minimal de perturbations qu'il faut pour atteindre la racine $h \in H$ en passant une et une seule fois par tous les états de $H \setminus \{h\}$, et donc exprime la vitesse à laquelle le numérateur de μ_h^ε tend vers 0 lorsque $\varepsilon \rightarrow 0$. Il suit que le rapport $\mu_h^\varepsilon / \mu_h^\varepsilon$ tend vers 0 (une constante $C_{h,h'} > 0$ respectivement) lorsque $\varepsilon \rightarrow 0$ si $v_h < v_{h'}$ ($v_h = v_{h'}$ respectivement). En conséquence, la distribution μ^* affecte des probabilités positives aux états $h \in H$ de potentiel stochastique minimal :

$$[h \in C(\mu^*)] \Leftrightarrow [v_h = \min_{h' \in H} v_{h'}] \quad (7)$$

La recherche des éléments de $C(\mu^*)$ se ramène ainsi à la recherche des arbres de recouvrement de résistance minimale. De plus, par (6) les candidats à la stabilité stochastique sont des états récurrents du processus non perturbé associé. Finalement, l'ensemble des états stochastiquement stables est l'ensemble des états récurrents du processus non perturbé de potentiel stochastique minimal. Les états stochastiquement stables sont les états visités le plus fréquemment par le système. Ceci ne signifie que les autres états ne seront jamais visités à long terme, mais que la fréquence des visites devient négligeable relativement à celle des états stochastiquement stables au fur et à mesure que le processus se déroule. Si le système possède plus d'un état stochastiquement stable, alors, à long terme, le processus effectuera des cycles entre ces états. Une interprétation économique possible est celle de cycles stochastiques de long terme : chaque état représente un régime régulier ponctué de déséquilibres conduisant le système vers un autre régime régulier.

V. — LE MODÈLE DE FICTITIOUS PLAY

Le modèle de Fictitious Play avec mémoire limitée et échantillonnage a été introduit par Young (1993), puis a été développé par Maruta (1997) et Young (1998). Young (1993) obtient un résultat de convergence du processus de Fictitious Play non perturbé pour les jeux finis à n joueurs faiblement acycliques. Un jeu est faiblement acyclique (Young, 1993 et repris dans Young,

1998) si chaque profil est connecté à un équilibre de Nash strict par un chemin de réponses optimales asynchrones. Young (1993 repris dans Young 1998, théorème 7.1) montre que chaque ensemble récurrent du processus de Fictitious Play non perturbé est constitué d'un seul état qui correspond à un équilibre de Nash strict du jeu à condition que le rapport entre la taille de l'échantillon et la taille de la mémoire, s/m , soit suffisamment petit. Le rôle de l'échantillonnage est de fournir une possible hétérogénéité des informations factuelles entre les joueurs : deux joueurs distincts peuvent échantillonner des profils différents dans l'histoire, construire des croyances non identiques, et ainsi réagir de manière différente face à une même histoire. De ce fait, les situations cycliques récurrentes ne sont pas favorisées. La faible acyclicité du jeu assure l'existence d'un chemin asynchrone conduisant à un équilibre de Nash strict, et la taille relative de l'échantillon rend possible la construction d'un tel chemin dans le modèle de Fictitious Play. Enfin, la répétition d'un équilibre Nash strict m fois définit une histoire tronquée absorbante. Si le jeu n'est pas faiblement acyclique, on ne peut pas caractériser les ensembles récurrents de ce processus de Fictitious Play sans hypothèses additionnelles sur les gains du jeu (cf. Young, 1998, théorème 7.2, et appendice p.167). Par contre, on peut énoncer un résultat de sélection pour le processus de Fictitious Play perturbé associé.

Proposition 2 - (Durieu, Solal, Tercieux, 2005)

Soit $p \leq n^{1-n}$ et soit Y l'ensemble p -optimal minimal du jeu. Si s/m est suffisamment petit, s suffisamment grand, les états stochastiquement stables du modèle de Fictitious Play avec mémoire limitée et échantillonnage sont les histoires tronquées composées de certains profils de Y .

Il y a plusieurs remarques à formuler. Premièrement, le résultat nous informe que seules les histoires tronquées composées de profils de Y peuvent être stochastiquement stables. Pour le cas particulier où $Y = \{x\}$ pour un $x \in X$, on conclut que l'histoire $h = (x, x, \dots, x)$ est l'unique histoire stochastiquement stable, autrement dit, l'équilibre de Nash p -dominant, $p \leq n^{1-n}$, correspond à l'unique état stochastiquement stable du processus de Fictitious Play avec mémoire limitée et échantillonnage. Deuxièmement, si $n = 2$, alors la borne sur p est $1/2$ et on conclut que si le jeu possède un équilibre de Nash $1/2$ -dominant, il correspond à l'unique état stochastiquement stable du processus de Fictitious Play avec mémoire limitée et échantillonnage. Sinon, les histoires stochastiquement stables sont construites à partir de l'ensemble $1/2$ -optimal minimal. Dans l'exemple 1, les histoires stochastiquement stables sont constituées de certains profils pris dans $\{1, 2, 3\} \times \{A, B, C\}$. On généralise ainsi le résultat obtenu par Young (1993) et Maruta (1997) pour les jeux à deux joueurs faiblement acycliques. La proposition 2 indique que la faible acyclicité du jeu n'est pas nécessaire pour obtenir le résultat de sélection stochastique. Troisièmement, même si le résultat de la proposition 2 ne permet pas toujours de caractériser $C(\mu^*)$, il reste néanmoins utile dans la mesure où il permet

d'évacuer certaines histoires candidates à la stabilité stochastique. L'exemple suivant emprunté à Young (1993) illustre ce point.

Exemple 7

Considérons le jeu à deux joueurs de matrice des gains :

	A	B	C
1	(6, 6)	(0, 5)	(0, 0)
2	(5, 0)	(7, 7)	(5, 5)
3	(0, 0)	(5, 5)	(8, 8)

$(1,A)$, $(2,B)$ and $(3,C)$ sont les équilibres de Nash stricts de ce jeu : $(1,A)$ est $7/8$ -dominant ; $(2,B)$ est $3/5$ -dominant, et $(3,C)$ est $5/8$ -dominant. Mais $\{2, 3\} \times \{B, C\}$ est l'unique ensemble $1/2$ -optimal minimal du jeu. Par la proposition 2, les histoires stochastiquement stables sont construites à partir de profils de stratégies pris dans $\{2, 3\} \times \{B, C\}$. En effet, Young (1993) montre que l'unique histoire stochastiquement stable est la répétition du profil de Nash $(2,B)$. Néanmoins, pour obtenir ce résultat, Young calcule le potentiel stochastique des histoires composées des profils $(1,A)$, puis des profils $(2,B)$, et enfin des profils $(3,C)$. Grâce à la proposition 2, nous pouvons d'emblée exclure l'histoire composée de profils $(1,A)$. Ainsi, la recherche de l'histoire de potentiel stochastique minimal s'en trouve facilitée.

La preuve de la proposition 2 repose sur la comparaison entre deux résistances : le nombre minimal r_1 de perturbations afin que le processus quitte Y comparé au nombre minimal r_2 de perturbations pour que le processus entre dans Y à partir de la situation la plus défavorable (en termes de nombre de perturbations). Lorsque $p \geq n^{1-n}$, alors $r_2 < r_1$, ce qui signifie qu'il est plus facile pour le processus d'entrer dans Y que d'en sortir. Donc, lorsque $\varepsilon \rightarrow 0$ la fraction de temps passée par le processus en dehors de Y devient négligeable relativement à la fraction de temps passée par le processus dans Y . Pour conclure, il faut retenir que le processus de Fictitious Play avec mémoire limitée et échantillonnage sélectionne les profils les moins risqués c'est-à-dire ceux qui résistent le mieux à l'échec d'implémentation des réponses optimales. Nous allons voir dans la section suivante que le processus d'imitation de Josephson et Matros (2004) sélectionne les profils optimaux au sens de Pareto.

VI. — LE MODÈLE D'IMITATION

La règle d'imitation du modèle de Josephson et Matros (2004) est celle de l'exemple 5. À la différence du modèle de Fictitious Play les joueurs ne forment pas (ou ne tiennent pas compte) des croyances sur les choix futurs de leurs opposants. Par contre, comme dans le modèle de Fictitious Play de Young (1993), ils ont une mémoire limitée et échantillonnent un sous-

ensemble de profils dans l'histoire tronquée à partir duquel ils font leur choix. Le premier résultat obtenu par les auteurs est que le processus d'imitation non perturbé converge vers des histoires tronquées $h = (x, x, \dots, x) \in H$ composées d'un unique profil $x \in X$, histoires que nous noterons h_x par la suite. En effet, pour chaque profil $x \in X$, chaque histoire tronquée h_x est absorbante car quel que soit l'échantillon tiré par un joueur $i \in N$, il observe nécessairement la répétition de la stratégie $x_i \in X_i$ et du gain $u_i(x_i, x_{-i})$. Donc, sa règle d'imitation non perturbée lui impose d'imiter son action courante x_i , et ainsi le processus est absorbé dans l'histoire tronquée h_x . Ensuite, dès qu'une histoire tronquée contient au moins deux profils qui diffèrent sur une composante au moins, il est toujours possible de construire une suite de transitions qui conduisent à une histoire tronquée composée d'un unique profil lorsque $2s < m$. L'intuition est que si un joueur tire l'échantillon le plus récent qui contient au moins deux stratégies distinctes, il imitera la plus profitable dans cet échantillon, et si ce type d'opportunité se répète s fois, alors il copiera s fois cette stratégie. Dans le même temps, les autres joueurs, n'ayant pas reçu l'opportunité d'ajuster leur stratégie, ont répété s fois la même stratégie. Donc, pour $m > 2s$, les s plus récentes périodes sont constituées d'un même profil de stratégies. Maintenant, il suffit que n'importe quel joueur ayant l'opportunité d'ajuster sa stratégie tire cet échantillon de s profils identiques pour qu'il recopie cette stratégie. Si ce processus se répète encore $m - s$ fois, on obtient une histoire tronquée composée d'un unique profil. En combinant ce raisonnement avec le point précédent, on arrive à la conclusion que les histoires tronquées composées d'un unique profil constituent les ensembles récurrents du processus d'imitation non perturbé. Par (6) les candidats à la stabilité stochastique sont les histoires tronquées de type $h_x, x \in X$. Le résultat suivant, à mettre en contraste avec la proposition 1, établit une relation entre les états stochastiquement stables et les ensembles fermés par des réponses améliorantes minimaux.

Proposition 3 - (Josephson, Matros, 2004, p. 251)

Le sous-ensemble de profils $Z = \{x \in X : h_x \in C(\mu^)\}$ est une réunion d'ensembles fermés par des réponses améliorantes minimaux.*

La preuve de cette proposition est construite en plusieurs étapes. Tout d'abord, (6) indique que les histoires $h_x, x \in X$, sont les seules histoires susceptibles d'être stochastiquement stables. On peut donc restreindre l'analyse à l'ensemble des arbres de racine $h_x, x \in X$, et se concentrer sur les résistances entre des paires de ces histoires. On notera $r(h_x, h_z)$ la résistance entre deux de ces histoires sur un de ces arbres, étant entendu que cette transition nécessite en réalité plusieurs périodes.

i) $r(h_x, h_z) \geq 1$ pour tout $z \neq x$ et vaut 1 si et seulement si $z \in B(x)$. Par définition, il faut au moins une perturbation pour transiter d'une histoire absorbante vers une autre, et une seule perturbation est requise si et seulement si z peut être atteint en partant de x via une réponse améliorante puisqu'il suffit au

joueur en question, disons i , d'implémenter par « erreur » (une perturbation) cette réponse améliorante et de la copier $m - 1$ fois de manière consécutive, ce qui est probable puisque $v_i > 0$.

ii) Si $z \in B(x)$, alors $v_{h_z} \leq v_{h_x}$. Considérons un arbre de recouvrement τ_{h_x} de résistance minimale i.e. $r(\tau_{h_x}) = v_{h_x}$. Construisons un arbre de recouvrement τ_{h_z} à partir de τ_{h_x} de la façon suivante : on ajoute la transition $h_x \rightarrow h_z$ et on supprime la transition $h_z \rightarrow h_y$, $y \in X$, sur l'arbre τ_{h_x} . On obtient bien un arbre de recouvrement $r(\tau_{h_z})$ de résistance $r(\tau_{h_z}) = v_{h_x} + r(h_x, h_z) - r(h_z, h_y)$, où $r(h_x, h_z) = 1$, et $r(h_z, h_y) \geq 1$ par le point i). Donc $v_{h_z} \leq v_{h_x}$.

iii) Pour $x, z \in X$ appartenant à un même ensemble fermé par réponses améliorantes minimal, on a $v_{h_z} = v_{h_x}$. Pour voir ceci, il suffit de remarquer qu'il existe un chemin (une suite de profils dont deux consécutifs diffèrent d'une seule composante) de réponses améliorantes de z vers x et vice versa par définition d'un tel ensemble fermé minimal. Par le point ii) on a $v_{h_z} \leq \dots \leq v_{h_x}$ et $v_{h_z} \geq \dots \geq v_{h_x}$, donc $v_{h_z} = v_{h_x}$.

iv) Si $z \in X$ n'appartient pas à un ensemble fermé par réponses améliorantes minimal, alors il existe $x \in X$ tel que $v_{h_z} > v_{h_x}$ (Josephson, Matros, 2004, lemme 5).

La proposition 3 découle des étapes i) –iv). De l'étape iv), on conclut qu'une histoire h_x , où $x \in X$ est un profil qui n'appartient pas à un ensemble fermé par réponses améliorantes minimal ne peut pas être stochastiquement stable. Du point iii), on sait que tous les profils $x \in X$ d'un même ensemble fermé par réponses améliorantes minimal génèrent des histoires h_x de même potentiel stochastique. Enfin, par (7) les histoires stochastiquement stables sont celles qui ont un potentiel stochastique minimal. La proposition 3 ne dispense pas du programme d'optimisation (7). Par contre, on sait repérer les histoires qui ne seront pas stochastiquement stables, et ainsi le programme s'en trouve allégé. Pour une certaine classe de jeux, Josephson et Matros identifient l'ensemble des histoires stochastiquement stables.

Proposition 4 - (Josephson, Matros, 2004, p. 252)

Si un ensemble fermé par des réponses améliorantes minimal $Z = \prod_{i \in N} Z_i \subseteq X$ est un produit cartésien de sous-ensembles de stratégies qui domine strictement au sens de Pareto toutes les stratégies en dehors de cet ensemble, alors il existe une taille de l'échantillon pour laquelle $\{h_z \in H : z \in Z\}$ est le sous-ensemble des histoires stochastiquement stables du processus d'imitation avec mémoire limitée et échantillonnage.

Pour bien saisir l'intuition de ce résultat considérons un exemple.

Exemple 8

Considérons le jeu symétrique à deux joueurs suivant :

	A	B
A	(2, 2)	(3, 0)
B	(0, 3)	(4, 4)

Les deux équilibres de Nash stricts (A,A) et (B,B) sont les deux ensembles fermés par des réponses améliorantes [optimales] minimaux du jeu. Par la proposition 3, $h_{(A,A)}$ et $h_{(B,B)}$ sont les deux histoires candidates à la stabilité stochastique. Cependant, (B,B) domine strictement au sens de Pareto (A,A) , (B,A) et (A,B) et ainsi on conclut que $h_{(B,B)}$ est l'unique histoire stochastiquement stable du processus d'imitation en appliquant la proposition 4. En effet, considérons $h_{(B,B)}$. Il est clair que deux perturbations successives ne peuvent pas faire transiter le processus de $h_{(B,B)}$ vers $h_{(A,A)}$, car les gains obtenus par les joueurs sur le profil (B,A) (ou (A,B)) puis le profil (A,A) sont inférieurs aux gains obtenus par les deux joueurs sur le profil (B,B) . En conséquence, les stratégies implémentées par erreur n'ont aucune chance d'être copiées par les joueurs aux périodes ultérieures quel que soit l'échantillon tiré (en supposant que $s > 1$). Par contre, la transition de l'histoire $h_{(A,A)}$ vers l'histoire $h_{(B,B)}$ peut se réaliser à l'aide de deux perturbations. Si les deux joueurs implémentent successivement par erreur la stratégie B , puis, lorsqu'ils en ont l'opportunité, tirent ensuite sur un nombre suffisant de périodes un échantillon de profils dans lequel (B,B) figure, ils copieront autant de fois la stratégie B puisqu'elle aura procuré le gain le plus élevé de l'échantillon, et ainsi l'histoire $h_{(B,B)}$ émergera à partir de l'histoire $h_{(A,A)}$. Il est maintenant facile de conclure que $vh_{(B,B)} < vh_{(A,A)}$. Par contre, on peut remarquer que dans le modèle de Fictitious Play avec mémoire limitée et échantillonnage, l'histoire $h_{(A,A)}$ est la seule histoire tronquée stochastiquement stable. En effet, les équilibres de Nash stricts (A,A) et (B,B) correspondent aux deux ensembles fermés par des réponses optimales minimaux et $\{(A,A)\}$ est 1/3-optimal (ou bien (A,A) est 1/3-dominant). La proposition 2 s'applique. Alors que le modèle d'imitation sélectionne l'issue optimale au sens de Pareto, le modèle de Fictitious Play sélectionne l'issue la moins risquée.

VII. — CONCLUSION

Nous avons pu constater que la contagion des stratégies est très sensible aux détails du processus d'adaptation. Elle est également très sensible au processus de perturbation. Nous n'avons pas orienté notre article dans cette direction, mais cette voie de recherche existe. Nous n'avons pas non plus abordé la question des applications économiques des modèles de Fictitious Play et d'imitation. Elles se développent également. L'article suivant de Billand et Bravard explore un processus de concurrence entre firmes dont le comportement est

basé sur l'imitation. Les auteurs centrent l'analyse sur une concurrence à la Cournot et, dans ce cadre, envisagent différentes procédures d'imitation perturbées dont celle proposée par Josephson et Matros (2004). Quoi qu'il en soit, on retiendra que les modèles de contagion évolutionnaires formalisent le concept d'ordre spontané. À ce titre, l'évolution des conventions, de normes et d'institutions organiques (la monnaie par exemple) peut être l'objet d'une formalisation à l'aide des éléments présentés dans cet article. Qualitativement, les processus d'évolution que nous avons décrits captent aussi bien les phénomènes de convergence vers des situations d'équilibre que les transitions entre ces équilibres. D'un côté, une série d'adaptations peut conduire le système vers une position d'équilibre. D'un autre côté, l'introduction de perturbations dans les règles de décision fragilise ces situations d'équilibre et provoque une série de désajustements qui amène le système dans une nouvelle position d'équilibre. Ces transitions, quoique peu fréquentes, évitent les cas de verrouillage à long terme sur les équilibres les moins stables du système. Autrement dit, les choix initiaux ne déterminent pas les états observés à long terme.

Voir références page suivante

RÉFÉRENCES

- BASU K., and WEIBULL J.-W. (1991), « Strategy Subsets Closed Under Rational Behavior », *Economics Letters* 36, 141-146.
- BARON R., DURIEU J., HALLER H., SOLAL P. (2002), « A Note on Control Costs and Logit Rules for Strategic Games », *Journal of Evolutionary Economics* 12, 563-575.
- BLUME L. (1993), « Statistical Mechanics of Strategic Interaction », *Games and Economic Behavior* 5, 387-426.
- DURIEU J., SOLAL P., TERCIEUX O. (2005), « Adaptive Play and Curb set Selection », mimeo.
- JOSEPHSON J., MATROS A. (2004), « Stochastic Imitation in Finite Games », *Games and Economic Behavior* 49, 244-259.
- MARUTA T. (1997), « On the Relationship between Risk-Dominance and Stochastic Stability », *Games and Economic Behavior* 19, 221-234.
- MORRIS S., ROB R., SHIN H.-S. (1995), « p -Dominance and Belief Potential », *Econometrica* 63, 145-157.
- ROBSON, A.-J., VEGA-REDONDO F. (1996), « Efficient Equilibrium Selection in Evolutionary Games with Random Matching », *Journal of Economic Theory* 70, 65-92.
- SOLBERG J. (1975), « A Graph Theoretic Formula for the Steady State Distribution of Finite Markov Processes », *Management Science*, 21 :1040-1048.
- TERCIEUX O. (2004), « p -Best Response Sets and the Robustness of Equilibria to Incomplete Information », *Games and Economic Behavior* forthcoming.
- WALLISER B. (1998), « A Spectrum of Equilibration Processes in Game Theory », *Journal of Evolutionary Economics* 8, 67-87.
- YOUNG H.-P. (1993), « The Evolution of Conventions », *Econometrica* 61, 57-84.
- YOUNG H.-P. (1998), « Individual Strategy and Social Structure. An Evolutionary Theory of Institutions ». Princeton : Princeton University Press.