De la stabilité évolutionniste à la stabilité stochastique. Réflexions sur les jeux évolutionnistes stochastiques

André Orléan

(Revue Economique, vol 47, n°3, mai1996, 589-600)

Abstract

Dans ses développements récents, la théorie des jeux évolutionnistes a modélisé l'existence de perturbations aléatoires affectant la dynamique de l'apprentissage. Ces travaux ont conduit au concept de stabilité stochastique, proposé par Foster et Young. Dans un jeu symètrique à deux stratégies, possèdant deux équilibres de Nash stricts, ce critère sélectionne celui qui satisfait au critère de risque-dominance de Harsanyi et Selten. Le présent article met en lumière certaines limites du critère de stabililité stochastique. Il souligne, en particulier, que si le processus stochastique considéré n'est pas ergodique, la stabilité stochastique n'est pas définie. L'article se termine sur une analyse de quelques exemples de jeux évolutionnistes stochastiques non ergodiques.

Résumé anglais :

Recently, evolutionary games has analyzed the way evolutionary dynamics is modified when the system is subjected to random perturbations. This approach has led to a new concept of stability, the one of stochastic stability proposed by Foster and Young. For 2×2 symmetric games with two symmetric strict Nash equilibria, there is only one stochastically stable equilibrium, the one which satisfies Harsanyi and Selten's criterion of risk dominance. This paper reviews the recent literature on this domain and highlights some limits of the stochastic stability. In particular, if the process is non-ergodic, the later criterion

cannot be applied. We finish by studying some non-ergodic stochastic evolutionary game dynamics.

1 Introduction

Une des raisons du succès rencontré auprès des économistes par les jeux évolutionnistes, provient du fait qu'ils fournissent à la théorie des jeux un cadre dynamique reposant sur des hypothèses comportementales simples. Celles-ci supposent que les agents déterminent leur choix en comparant les résultats passés des différentes stratégies et en privilégiant celles dont les gains sont les meilleurs. Un vaste ensemble de spécifications est compatible avec ces hypothèses générales. Si l'on fait l'hypothèse qu'il existe un ensemble fini de stratégies $i = 1, 2, ..., k$ et qu'on note $\mathbf{n}(t) = [n_1(t), ..., n_k(t)]$, le vecteur formé à partir des nombres d'individus ayant choisi, à la date t, les k différentes stratégies, il découle des hypothèses précédentes que la dynamique suivie par $\mathbf{n}(t)$ est telle que :

$$\frac{n_i(t+1) - n_i(t)}{n_i(t)} > \frac{n_j(t+1) - n_j(t)}{n_j(t)} \quad \text{si et seulement si} \quad u_i\left[\mathbf{n}(t)\right] > u_j\left[\mathbf{n}(t)\right] \tag{1}$$

où $u_h(\mathbf{n})$ désigne le gain obtenu lorsque la variable d'état vaut $\mathbf{n}$ et qu'on a choisi la stratégie h. Telle est la transposition à l'économie du principe darwinien d'évolution. Notons que la logique d'apprentissage ainsi spécifiée repose sur une rationalité des agents limitée. Ils sont myopes et relativement inertes[1] : ils sont myopes dans la mesure où ils ne prennent pas en compte le fait qu'en $(t+1)$, les autres joueurs auront également modifié leur stratégie, ce qui influera sur l'utilité qu'engendrera alors la stratégie qu'ils auront choisie ; ils sont également inertes dans la mesure où ils peuvent conserver, pendant un certain temps, un choix qui procure des gains médiocres avant d'en changer pour une meilleure stratégie[2].

[1] Ce caractère relatif dépend de la spécification retenue. Seule une inertie totale est incompatible avec les hypothèses précédentes ; mais on peut, par exemple, avoir une absence totale d'inertie (voir, plus loin, la dynamique dite de meilleure réponse dans [[5]]).

[2] Il n'entre pas dans le cadre de cet article de discuter la pertinence de ces hypothèses. Notons simplement que Kandori, Mailath et Rob [1993] les justifient en montrant qu'on peut les considérer comme découlant d'une conduite rationnelle.

L'intérêt de disposer d'un cadre dynamique pour analyser un jeu est lié aux difficultés rencontrées dans l'utilisation du concept d'équilibre de Nash (EN). Si l'on sait qu'à l'équilibre de Nash, de par sa définition même, aucun joueur n'est incité à modifier unilatéralement son choix, la théorie des jeux ne nous dit rien sur la manière dont on parvient à cet équilibre, ce qui s'avère particulièrement gênant quand le jeu considéré possède plusieurs EN. Le cadre évolutionniste fournit une réponse précise à cette question puisqu'il spécifie la dynamique du système. Les premiers travaux en jeux évolutionnistes ont raisonné dans un cadre uniquement déterministe. Ils ont conduit à une nouvelle notion d'équilibre, à savoir les points fixes stables de la dynamique évolutionniste (PFS). Le critère de stabilité dynamique est proche de celui proposé par Maynard Smith dans ses travaux pionniers [6] : la stabilité évolutionniste. La stratégie i est une stratégie stable évolutivement (SES) si, partant d'une situation où toute la population l'a adoptée, l'apparition soudaine d'un groupe de taille ϵ suffisamment petite ayant adopté une stratégie j, dite stratégie mutante, ne réussit pas à envahir la population, et cela quel que soit j différent de i. La stabilité dynamique ou la stabilité évolutionniste permet-elle de sélectionner au sein de l'ensemble des EN, une catégorie significativement plus étroite et robuste ? La réponse est malheureusement négative. Sans entrer dans le détail fastidieux de la comparaison précise des trois sortes d'équilibre, EN, PFS et SES, (se reporter à [4]), il nous suffira de noter que tous les équilibres de Nash stricts (ENS) sont, à la fois, des PFS et des SES. Un peu de recul permet de comprendre aisément ce résultat : la définition des trois types d'équilibre repose uniquement sur des propriétés locales. Autrement dit, les deux critères de stabilité n'exploitent que très peu le cadre dynamique que proposent les jeux évolutionnistes ; seul ce qui se passe au voisinage des équilibres y est pris en compte.

Dans le même temps où ces résultats émergeaient, il est apparu aux théoriciens des jeux évolutionnistes que la seule dynamique d'apprentissage ne modélisait qu'une dimension partielle de la réalité, dans la mesure où de nombreuses perturbations, en particulier de nature stochastique, affectent simultanément l'évolution de la population. Ce fait est d'ailleurs à l'origine même du concept de SES puisqu'il s'agissait pour Maynard Smith de faire valoir que seules les stratégies capables de résister à des chocs mutationnels constituaient des équilibres possibles de la population. Cependant Maynard Smith, pour définir une SES, ne prend en compte que des chocs mutationnels *ponctuels et isolés* : un groupe ϵ de mutants apparaît soudainement, puis le système est laissé à lui-même jusqu'au prochain choc. Comme le choc est de faible amplitude, le système demeure dans le voisinage de la position initiale ; et comme ce choc est isolé, c'est la seule dynamique d'apprentissage qui guide ensuite le système. Il s'ensuit que, si la position initiale est un PFS, le système reviendra à son point de départ et ne gardera nulle trace du choc. Foster et Young (FY) [2] notent que, si on introduit des chocs stochastiques de faible amplitude mais continuels, la question devient plus complexe. Que le système retourne alors vers l'équilibre de départ n'a rien de trivial car la variable d'état peut avoir été conduite, sous l'effet de la série des chocs aléatoires, hors du voisinage de son point de départ. Il en découle que les caractéristiques de la dynamique globale deviennent alors essentielles pour comprendre l'évolution du système. La prise en compte de perturbations stochastiques permanentes débouche alors sur un nouveau critère, la stabilité stochastique, proposé par FY. Le but de cet article est de présenter aux lecteurs ce nouveau concept ; de montrer que, contrairement aux critères de stabilité précédents, il permet de discriminer efficacement entre les EN ; et enfin d'en souligner certaines limites.

2 Le modèle de base

Pour mener à bien ces tâches, on s'appuiera sur des analyses qui ne traiteront que des jeux symètriques à deux stratégies, notées s_1 et s_2, tels que :

$$\mathrm{V} = \begin{array}{c c} & \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{array}{c} s_1 \\ s_2 \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array} \quad \text{vérifiant} \begin{cases} a > c \quad \text{et} \quad d > b \\ a < d \\ a + b > c + d \end{cases} \tag{2}$$

Les conditions (2) impliquent que le jeu V possède deux ENS en stratégies pures : (s_2, s_2) qui est pareto-optimal et (s_1, s_1) qui risque-domine (s_2, s_2). On notera N, le nombre total de joueurs. Puisque le nombre de stratégies, k, vaut deux, la variable d'état $\mathbf{n}(t) = [n_1(t), n_2(t)]$ est parfaitement déterminée par la seule donnée de $n_1(t)$. On peut donc simplifier les notations générales en renommant la variable $n_1(t)$, $n(t)$. $p(t)$ désignera $n(t)/N$. Pour définir la dynamique, conformément à la démarche proposée par Maynard Smith, on suppose qu'à chaque période, les N individus se rencontrent deux par deux de manière aléatoire et jouent le jeu V. Si N est grand et si le nombre d'appariements aléatoires est suffisamment élevé, on a :

$$\begin{cases} u_1(p) = p \cdot a + (1-p) \cdot b \\ u_2(p) = p \cdot c + (1-p) \cdot d \\ u(p) = p \cdot u_1(p) + (1-p) \cdot u_2(p) \end{cases} \tag{3}$$

où $u_1(p)$, $u_2(p)$ et $u(p)$ désignent respectivement l'utilité obtenue par les agents ayant choisi la stratégie s_1, l'utilité obtenue par ceux ayant choisi s_2 et l'utilité obtenue en moyenne par la population. Parmi les dynamiques satisfaisant à la condition (1), il en est une privilégiée par les biologistes, et très souvent reprise par les économistes : la dynamique de réplication. Dans le cas d'un jeu à deux stratégies, elle s'écrit, en temps discret :

$$p(t+1) = p(t) \cdot \frac{u_1\left(p(t)\right)}{u\left(p(t)\right)} \quad \Leftrightarrow p(t+1) - p(t) = p(t)\left[1-p(t)\right] \frac{u_1\left(p(t)\right) - u_2\left(p(t)\right)}{u\left(p(t)\right)} \tag{4}$$

En temps continu et pour le jeu V, on obtient :

$$\begin{cases} \dfrac{dp(t)}{dt} = p(t)\left[1-p(t)\right]\left[u_1(p) - u_2(p)\right] = C\, p(t)\ \left[1-p(t)\right]\left[p(t) - \mu\right] = r\left[p(t)\right] \\ 0 < \mu = \dfrac{d-b}{(a-c)+(d-b)} < 1 \quad \text{et} \quad C = (a-c)+(d-b) > 0 \end{cases} \tag{5}$$

On montre aisément que la dynamique (5) a deux PFS : l'état $\{\, p=0\}$; et l'état $\{\, p=1\}$. Ce résultat n'est pas étonnant puisque les EN correspondant sont des ENS. La dynamique a également un point fixe instable, $\{p=\mu\}$, qui correspond à l'équilibre de Nash du jeu en stratégies mixtes. Si le critère de stabilité dynamique ne permet de discriminer entre les deux ENS, il est cependant possible de déterminer leur bassin d'attraction :

$$\begin{cases} \text{si } p(0) \in [0\ ,\ \mu[& p(t) \to 0 \\ \text{si } p(0) \in]\mu\ ,\ 1] & p(t) \to 1 \end{cases}$$

Remarquons que le bassin d'attraction de l'état 1 a une taille plus grande que celui de l'état 0. Cette propriété est équivalente à celle de risque-dominance. En effet μ est inférieur à 0.5 si et seulement si $(a-c) > (d-b)$.

Cette dynamique de base laisse de côté un aspect important de la logique évolutionniste, à savoir la présence constante de facteurs mutationnels. Les économistes désignent par "mutations", le fait qu'à chaque instant, une certaine proportion d'agents passe d'une stratégie à une autre, indépendamment des utilités obtenues, sur la seule base de croyances idiosyncrasiques concernant les qualités supposées des stratégies. On peut penser ce processus comme étant un simple phénomène de "main tremblante". Les mutations constituent une source autonome et permanente de variabilité. Elles crèent un "bruit" perpétuel dans l'évolution. Avant d'en proposer une modélisation stochastique dans la section suivante, nous en analyserons les effets sur la détermination des équilibres dans le cadre d'un modèle qui reste déterministe.

On supposera qu'à chaque instant, il y a une proportion m_{12} de joueurs qui passe de la stratégie s_1 à la stratégie s_2, et une proportion m_{21} qui fait l'inverse. Les paramètres m_{ij} sont appelés des taux de mutation. Lorsque ces deux taux sont égaux à m, on obtient :

$$\frac{dp(t)}{dt} = C \cdot p(t)\,[1 - p(t)]\,[p(t) - \mu] + m\,[1 - 2\,p(t)] = h\,[p(t)\,,\,m] \qquad (6)$$

Notons que la dynamique reste déterministe, mais alors que les situations d'unanimité étaient des points fixes de la dynamique de réplication simple (5), cela n'est plus vrai pour la dynamique (6) qui repousse les points fixes vers l'intérieur de l'intervalle $[0, 1]$. Les états d'équilibres correspondent alors à des populations hétérogènes. C'est la permanence des mutations qui produit constamment de l'hétérogénéité, là où la dynamique de réplication tend à l'uniformité.

Etudions la fonction $h(p, m)$ de l'équation (6). Quand m croît de 0 à m^*, les deux PFS de la dynamique sont repoussés vers l'intérieur du segment $[0, 1]$: la taille des minorités non conformistes présentes dans les deux équilibres croît. Pour m égal à m^*, la dynamique globale se modifie radicalement. Le diagramme des phases est affecté par un bifurcation : le PFS correspondant à la stratégie paréto-dominante s_2 disparait ; seul demeure le PFS où est majoritaire la convention s_1. L'introduction des "mutations" endogènes a donc des effets importants sur la détermination des attracteurs. Plus m croît, plus le bassin d'attraction de la stratégie risque-dominante augmente, jusqu'à remplir tout l'espace des états. L'explication en est simple : par définition, la stratégie risque-dominante est celle qui s'adapte le mieux à un environnement hétérogène ; alors que l'efficacité de la stratégie parétooptimale ne se révèle pleinement que dans des contextes de forte homogénéité des choix. Aussi, lorsque l'hétérogénéité des PFS augmente avec m, l'utilité qu'y obtient la stratégie s_2, diminue relativement à celle que procure s_1. Pour m suffisamment grand, à savoir m^*, l'utilité liée à s_2 devient inférieure à celle qu'engendre s_1, et le PFS correspondant disparaît. Que deviennent ces résultats dans un cadre probabiliste ?

3 Définition de la stabilité stochastique

La recherche de Kandori, Mailath et Rob (KMR) [5] porte précisément sur la construction d'une modélisation stochastique des mutations. Par rapport au modèle de base présenté dans les sections précédentes, KMR ne supposent plus que N soit nécessairement très élevé. Par ailleurs ils raisonnent en temps discret et prennent n comme variable d'état plutôt que p. Dans ces conditions, les équations (3) s'écrivent :

$$\left\{ \begin{array}{l} u_1(n) = \dfrac{n-1}{N-1}\,a + \dfrac{N-n}{N-1}\,b \quad \text{pour} \;\; 1 \le n \le N \\ u_2(n) = \dfrac{n}{N-1}\,c + \dfrac{N-n-1}{N-1}\,d \quad \text{pour} \;\; 0 \le n \le N-1 \end{array} \right. \tag{7}$$

On peut interpréter ces utilités à l'aide de l'hypothèse d'un grand nombre d'appariements aléatoires, comme précédemment, ou supposer qu'elles sont le résultat d'un processus de rencontres qui est tel qu'à chaque période, chaque agent rencontre les $(N-1)$ autres agents. Une des forces de leur analyse est de ne pas se centrer uniquement sur la dynamique de réplication, mais de raisonner sur l'ensemble des dynamiques vérifiant (1). Si l'on note :

$$n(t+1) = b\,[n(t)]$$

on se persuade facilement que, dans le cas de deux stratégies, la condition (1) est équivalente à :

$$sign\,[b(n)-n] = sign\,[u_1(n)-u_2(n)] \qquad \text{pour} \;\; n \neq 0, N \tag{8}$$

Un cas particulier est celui de la dynamique dite de meilleure réponse, qui suppose qu'à chaque instant, tous les agents se portent sur la stratégie ayant le mieux réussi à l'instant précédent, soit, en la notant $B(n)$:

$$B(n) = \left\{ \begin{array}{lcc} N & \text{si} & u_1(n) > u_2(n) \\ n & \text{si} & u_1(n) = u_2(n) \\ 0 & \text{si} & u_1(n) < u_2(n) \end{array} \right\}$$

Les PFS de ces dynamiques[3] sont, comme précédemment, les états d'unanimité $\{n = N\}$ et $\{n = 0\}$. La définition des deux bassins d'attraction est indépendante de la forme spécifique que prend $b(n)$.

KMR introduisent alors la possibilité de mutations aléatoires. Chaque individu a une probabilité ϵ de changer son choix. Le nombre d'individus qui, ayant choisi s_2, se reporte en définitive sur s_1, est noté $x(t)$. Les "mutations" individuelles étant indépendantes, $x(t)$ suit une loi binomiale. Il en est de même pour $y(t)$, le nombre d'individus qui mutent vers s_2. Il vient :

$$n(t+1) = b\left(n(t)\right) + x(t) - y(t) \qquad \text{avec} \begin{cases} x(t) \sim Bin\left(N - b\left(n(t)\right)\ ,\ \epsilon\right) \\ y(t) \sim Bin\left(b\left(n(t)\right)\ ,\ \epsilon\right) \end{cases} \tag{9}$$

L'équation (9) définit une chaîne de Markov sur l'espace des $(N+1)$ états, $Z = \{0, 1, ..., N\}$. Il est facile de montrer que cette chaîne est ergodique : elle admet une unique distribution stationnaire que l'on notera $q(\epsilon)$. On a $q(\epsilon) = (q(\epsilon)_0, q(\epsilon)_1, ..., q(\epsilon)_N)$ où $q(\epsilon)_n$ est la probabilité de l'état n. Cette probabilité s'interprète comme la proportion de temps que le groupe passe dans l'état n (quand t tend vers l'infini). On a $q(\epsilon)_n > 0$ pour tous les états. Cette proposition semble nous éloigner de la possibilité de sélectionner un seul état d'équilibre puisque désormais, quand t tend vers l'infini, le système visite perpétuellement les $(N + 1)$ éléments de l'espace Z des états. Cela provient du fait que les mutations rendent possible toutes les transitions d'un état i à un état j. Par exemple, si on considère la dynamique de meilleure réponse, et si on suppose que $n(t) = N$, on a $b\left(n(t)\right) = B(N) = N$. Si seule cette dynamique déterministe était prise en compte, le système resterait en N. Cependant, dès lors qu'on introduit des mutations, on peut peut obtenir, en $(t + 1)$, tous les états. En particulier, on observe $n(t + 1) = 0$ avec la probabilité $\epsilon^N > 0$.

L'idée de KMR est d'étudier comment se comporte la distribution stationnaire $q(\epsilon)$ quand ϵ tend vers zéro. Lorsqu'il existe une distribution-limite $q^* = \lim_{\epsilon \to 0} q(\epsilon)$, KMR proposent de définir les équilibres de long terme comme les états $n \in Z$, tels que $q^*_n > 0$. Nous verrons que cette définition est un cas particulier d'un critère plus général, celui de stabilité stochastique. KMR démontrent alors que, pour le jeu V, il n'existe qu'un seul équilibre

[3] La détermination de $u_i(n)$ et de $b(n)$ pour les états 0 et N admet plusieurs formulations alternatives, conformes à l'intuition ; mais nous ne rentrerons pas dans le détail des démonstrations.

de long terme, l'équilibre risque-dominant, à savoir l'état $\{n = N\}$. *Ce résultat est vrai quelle que soit la forme de la dynamique déterministe* $b(n)$ considérée. Il découle uniquement du fait que, quand s_1 risque-domine s_2, le bassin d'attraction de $\{n = N\}$ est plus grand que celui de $\{n = 0\}$. L'intuition derrière ce résultat est la suivante. Supposons que $N = 6$ et que le bassin d'attraction de la convention s_2 soit $(0, 1, 2)$; celui de la convention s_1 étant $(3, 4, 5, 6)$, alors, pour passer de l'état 6 au bassin d'attraction de l'état 0, il faut au minimum 4 mutations. La probabilité d'un tel événement est de l'ordre de ϵ^4. Pour passer de l'état 0 au bassin d'attraction de l'état 6, il faut au minimum 3 mutations, événement dont la probabilité est de l'ordre de ϵ^3. Quand $\epsilon \to 0$, le rapport de ces deux probabilités tend vers ∞ : il devient "infiniment" plus probable de passer de 0 à 6 que l'inverse. Pour cette raison, la probabilité se concentrera sur l'état 6.

3.1 Le modèle de Foster et Young

Dans un article important, Foster et Young (FY) [2] se posent la même question que celle étudiée précédemment : que se passe-t-il lorsqu'on introduit dans un modèle de jeux évolutionnistes, des perturbations aléatoires ? A la différence du modèle KMR, FY n'explicitent pas précisément la source de ces aléas. Il est fait référence à des facteurs environnementaux, aux aléas des rencontres aléatoires qui sont à la base des formules (3) de calcul des utilités, et à l'existence de mutations. Par ailleurs, ils ne s'intéressent qu'à la dynamique de réplication et suppose un temps continu. L'innovation fondamentale qu'ils introduisent, est d'utiliser une équation différentielle stochastique pour modéliser le rôle de ces perturbations. L'équation qu'ils étudient est obtenue en ajoutant à la dynamique de réplication (5), un processus de Wiener standard. On obtient :

$$dp(t) = C \cdot p(t) \cdot [1 - p(t)]\,[p(t) - \mu]\; dt + \sigma\, dW(t) = r\,[p(t)] + \sigma\, dW(t) \quad (10)$$

où W(t) est un processus de Wiener standard. Le fait d'introduire ce processus de manière additive pose immédiatement un problème puisqu'alors rien n'assure que $p(t)$ restera compris entre 0 et 1. Ceci conduit FY à restreindre l'espace des états à $[\Delta\,,\ 1 - \Delta]$, avec Δ suffisamment petit, et à supposer que les états extrêmes, Δ et $1 - \Delta$, sont des barrières réfléchissantes. FY démontrent alors que le processus ainsi défini est ergodique. Il admet une unique distribution limite, notée $f_\sigma(p)$, indépendante de la situation initiale et dont le support est $[\Delta\,,\ 1-\Delta]$. Lorsque p se trouve dans le voisinage de $1-\Delta$, les forces de sélection $r(p)$ sont positives et poussent le système vers $1-\Delta$; mais, si la perturbation probabiliste prend une valeur w suffisamment négative pour que $\delta\,p = r(p) + w$ (équation 10) soit négatif, p va alors se diriger vers l'état Δ. On pourra alors observer une transition du voisinage de $1-\Delta$ à celui de Δ. La différence essentielle avec le modèle KMR, qui est une conséquence du choix d'un temps continu, est le fait que désormais la probabilité d'une telle transition ne dépend plus seulement de la taille des bassins d'attraction, mais également des valeurs prises par $r(p)$; elle n'est plus indépendante de la partie déterministe de la dynamique.

Lorsque σ diminue, remonter contre le courant $r(p)$ devient plus difficile, car la probabilité d'obtenir un choc w d'amplitude suffisante pour compenser l'intensité sélectionniste, décroît. Autrement dit le temps que prennent les transitions de la région (Δ) à $(1-\Delta)$, qu'on notera T_+, et de la région $(1-\Delta)$ à (Δ), qu'on notera T_-, augmentent quand σ diminue. Cependant, pour $\sigma > 0$, ils restent finis, les deux bassins communiquent entre eux et le processus reste ergodique. Analysant les déformations de $f_\sigma(p)$ quand σ diminue, FY soulignent que la masse de probabilité se concentre de plus en plus dans la région $(1-\Delta)$. Autrement dit, si T_+ et T_- sont tous deux des fonctions décroissantes en σ, $T_+(\sigma)$ décroît plus vite que $T_-(\sigma)$. Dans un sens, la transition devient plus difficile que dans l'autre, ce qui explique qu'au cours du temps, le système se trouvera plus fréquement dans le voisinage de $(1-\Delta)$. L'idée centrale de FY est alors d'étudier ce qui se passe quand $\sigma \to 0$, ce qui va les conduire à la notion de stabilité stochastique dont ils donnent la définition suivante ([2] p. 228) : l'état $\mathbf{p}^*$ est stochastiquement stable si, quand $\sigma \to 0$, tous les voisinages de $\mathbf{p}^*$ ont une probabilité positive, soit :

$$\forall \epsilon > 0 \ \overline{\lim_{\sigma\to 0}} \int\limits_{N_\epsilon(p^*)} f_\sigma(\mathbf{p})\, d\mathbf{p} > \mathbf{0} \quad \text{avec} \quad N_\epsilon(\mathbf{p}^*) = \{\mathbf{p} :\ |\mathbf{p}-\mathbf{p}^*| < \epsilon\}$$

Dans le cas d'un espace d'états fini, on retrouve les équilibres de long terme de KMR. Dans le cas du jeu V et de la dynamique (10), FY démontrent que le processus à un unique état stochastiquement stable : l'état $1-\Delta$, celui où presque tous les joueurs ont opté pour la stratégie risque-dominante s_1.

3.2 Sur quelques limites du critère de stabilité stochastique

Le critère de stabilité stochastique, dans le cas discret (modèle KMR) ou continu (modèle FY), permet donc une sélection efficace parmi les EN, à savoir l'EN risque-dominant. Il nous semble cependant que la pertinence empirique de ce critère peut être contestée. Nous ferons trois remarques. Premièrement, la grandeur des perturbations n'étant pas un paramètre de contrôle de l'économiste, il est difficile de savoir à quoi correspond, dans la réalité, le fait que ϵ tend vers 0. Deuxièment, même si nous supposons une situation où le bruit tend vers 0, il n'est pas clair que c'est l'équilibre stochastiquement stable qui sera effectivement observé. Pour le comprendre, considérons la chaîne de Markov à deux états, notés A et B, définie de la manière suivante :

$$\mathbf{P}_\epsilon = \begin{bmatrix} 1-\epsilon^2 & \epsilon^2 \\ \epsilon & 1-\epsilon \end{bmatrix} \quad \text{avec} \quad 0 < \epsilon < 1$$

$\mathbf{P}_\epsilon$ est ergodique : sa probabilité stationnaire est unique. On la notera π_ϵ. Elle est déterminée par la relation suivante :

$$\frac{\pi_\epsilon(A)}{\pi_\epsilon(B)} = \frac{p_{BA}}{p_{AB}} = \frac{\epsilon}{\epsilon^2} = \frac{1}{\epsilon} \tag{11}$$

La difficulté de l'étude du comportement de cette chaîne, lorsque $\epsilon \to 0$, vient du fait que $\mathbf{P}_0$ n'est pas ergodique et possède deux distributions stationnaires : π_A telle $\pi_A(A) = 1$ et π_B telle $\pi_B(B) = 1$. Cependant, lorsque $\epsilon \to 0$, on a, d'après l'équation (11), $\pi_\epsilon \to \pi_A$. Seul l'état A est stochastiquement stable. On peut mieux comprendre ce phénomène en examinant le temps de première transition d'un état à l'autre. Il s'agit d'une variable aléatoire dont on peut calculer aisément l'espérance. On obtient :

$$E(T_{A\to B}) = \frac{1}{p_{AB}} = \frac{1}{\epsilon^2} \quad \text{et} \quad E(T_{B\to A}) = \frac{1}{p_{BA}} = \frac{1}{\epsilon}$$

On retrouve qualitativement les situations précédemment décrites. Quand ϵ décroît, les transitions d'un état à l'autre deviennent de plus en plus difficiles. $E(T_{i\to j})$ tend vers l'infini, mais plus rapidement dans un sens que dans l'autre. Dans notre exemple, la transition de A vers B est la plus difficile de telle sorte que la probabilité, $\pi_\epsilon(A)$, se concentre sur A. Si on observe le processus durant un temps infini, le système demeurera en A une "infinité" de fois plus longtemps qu'en B.

Cependant les conséquences empiriques de cette propriété ne sont pas claires. En effet, supposons que l'état initial soit B, quand ϵ devient petit, $E(T_{B\to A})$ devient très grand, de telle sorte qu'en temps fini, même très long, la probabilité est très grande de ne pas observer de transition vers A. Comme le temps Θ d'observation est nécessairement fini, il existe une valeur de ϵ telle que $E(T_{B\to A})$ soit rendu beaucoup plus grand que Θ. Dans ces conditions, bien que A soit stochastiquement stable, la probabilité d'observer son apparition est très faible et peut être rendue aussi petite que l'on veut.

Une troisième limitation du critère de stabilité stochastique provient du fait qu'il n'est applicable que dans le cas d'un processus ergodique. Si, pour une valeur finie du bruit, qu'il soit mesurée par σ ou ϵ, il existe plusieurs distributions stationnaires, on ne peut plus appliquer ce critère. Comme le montreront les travaux de Fuddenberg et Harris puis ceux de Boyer et Orléan, cette dernière limitation peut s'avérer importante dans la mesure où l'ergodicité n'est en rien une propriété générale des jeux évolutionnistes stochastiques. Aussi convient-il de développer de nouveaux outils pour les analyser.

4 Jeux évolutionnistes stochastiques non ergodiques

Fuddenberg et Harris (FH) [3] critiquent l'approche additive de FY pour deux types de raison. D'une part, comme on l'a vu, l'approche additive les oblige à supposer l'existence de barrières réfléchissantes, hypothèse ad hoc qu'ils ne justifient pas. D'autre part, cette approche ne spécifie pas précisément la nature du choc. Par exemple, est-il plausible que la variance induite par ce choc sur la variable n soit constante que n soit très petit ou très grand ? Pour éviter ces écueils, FH proposent d'introduire le terme de perturbations dans l'équation qui décrit l'évolution de la taille des diverses sous-populations. Si l'on se place dans le cas général d'un jeu à k stratégies, ils obtiennent :

$$dn_i(t) = n_i(t) \cdot [u_i(\mathbf{n}(t))\, dt + \sigma_i\, dW_i(t)] \tag{12}$$

où $\mathbf{W} = (W_i)_{i=1,\ldots,k}$ est un processus de Wiener à k dimensions dont la matrice de variance-covariance est égale à l'identité. Il s'agit d'une série de k chocs indépendants, spécifiques à chaque stratégie, qui affectent globalement la formation des gains. Grâce au lemme d'Ito, FH obtiennent l'équation différentielle stochastique que vérifient les proportions. Dans le cas du jeu V, il vient :

$$dp = p\,(1-p)\left[(u_1(p) - u_2(p))\,dt + \left(\sigma_2^2\ (1-p) - \sigma_1^2\ p\right)dt + \sigma_1\,dW_1 - \sigma_2\,dW_2\right] \tag{13}$$

On observe trois termes. Le premier nous est familier. Il s'agit du terme associé à la dynamique de réplication (5). Le deuxième terme est nouveau. On y voit les chocs aléatoires modifier la partie déterministe de la dynamique. Le troisième terme modélise l'effet direct des chocs. Notons deux caractéristiques qui différencient fortement ce modèle de celui proposé par FY. D'une part, comme n_i est en facteur dans l'équation (12), il vient que la variance du processus de Wiener que suit cette variable, n'est pas constante : elle décroît quand n_i diminue. D'autre part, le terme $p(1-p)$ étant en facteur dans l'équation (13), les états $\{p = 0\}$ et $\{p = 1\}$ sont des points fixes de ce processus, quelles que soient les valeurs des σ_i.

Dans le cas du jeu V, FH obtiennent les résultats suivants. Il n'y a pas ergodicité : si σ_1 et σ_2 sont suffisamment petits, $p(t)$ converge vers un des deux états $\{p = 0\}$ et $\{p = 1\}$ avec une probabilité égale à 1. Mais, à la différence de la dynamique déterministe associée, quel que soit l'état initial, il existe une probabilité positive d'atteindre l'un quelconque de ces deux états. Cependant, quand σ_1 et σ_2 tendent vers 0, on se rapproche du cas déterministe : la probabilité que $p(t)$ converge vers $\{p = 1\}$, si $p(0) > \mu$, tend vers 1. Dans ce modèle, la modélisation probabiliste n'apporte guère plus d'informations que le modèle déterministe associé.

Dans un second temps, FH prennent en compte l'existence de taux de mutation déterministe. L'équation (12) devient dans le cas de deux stratégies i et j :

$$dn_i(t) = n_i(t) \cdot [u_i\,(\mathbf{n}(t))\,dt + \sigma_i\,dW_i(t)] - m_{ij}\,n_i(t)\,dt + m_{ji}\,n_j(t)\,dt$$

Utilisant le lemme d'Ito comme précédemment, on en déduit l'équation différentielle que vérifie p. Dans le cas du jeu V, le processus est ergodique $\forall \sigma_1\,,\ \sigma_2\,,\ m_{12}\,,\ m_{21} > 0$. Lorsque $\sigma_1\,,\ \sigma_2\,,\ m_{12}\,,\ m_{21}$ tendent vers 0, de telle sorte que $m_{12}\,/\,m_{21}$

reste constant, alors la distribution ergodique converge sur l'équilibre risque-dominant.

4.1 Dilemme du prisonnier itéré avec longueur aléatoire

Dans le modèle que proposent Boyer et Orléan (BO) [1], c'est la matrice V qui est aléatoire. Dans un premier temps, ils introduisent un dilemme du prisonnier classique dont les *pay-off*, en utilisant les notations habituelles, valent : $T > R > P > S$. Puis ils considèrent ce dilemme du prisonnier, répété l fois, et analysent la compétition entre la stratégie "Donnant, donnant", notée TFT en référence à sa dénomination anglaise *Tit-For-Tat*, et la stratégie consistant à faire défection systématiquement, notée ALLD. On établit facilement que la matrice du jeu ainsi défini s'écrit :

$$\mathrm{DPR} = \begin{array}{c c} & \begin{array}{cc} ALLD \qquad\qquad & \qquad TFT \end{array} \\ \begin{array}{c} ALLD \\ TFT \end{array} & \begin{bmatrix} P \cdot l & T + P \cdot (l-1) \\ S + P \cdot (l-1) & R \cdot l \end{bmatrix} \end{array}$$

Cette matrice est aléatoire car la longueur l du jeu est supposée aléatoire. On fait en effet l'hypothèse que chaque coup peut être le dernier coup avec la probabilité s. On a donc :

$$\Pr ob(l = n+1) = s \cdot (1-s)^n \qquad \Rightarrow E(l) = L = \frac{1}{s} \tag{14}$$

L'analyse évolutionniste traditionnelle consiste à analyser le jeu DPR en remplaçant l par son espérance L. TFT est alors un ENS paréto-dominant si $L > L_0 = \frac{T-P}{R-P}$, et $ALLD$ est risque-dominant si $L < \frac{T}{R-P}$. La dynamique de réplication associée converge, en fonction de la situation initiale, vers l'une ou l'autre des situations d'unanimité : $\{p = 0\}$ ou $\{p = 1\}$, où p est la proportion des TFT. BO proposent de reprendre cette question en intégrant le fait qu'à chaque round t, la longueur du jeu ne vaut pas L, mais $l(t)$, longueur tirée selon la loi géomètrique (14). Raisonnant en temps discret, la dynamique de réplication (4) pour le jeu DPR s'écrit, en supposant $S = 0$:

$$p(t+1) - p(t) = p(t)\,(1 - p(t))\,\frac{p(t)\,[(R-P)\,l(t) + 2P - T] - P}{u\,(p(t), l(t))} \tag{15}$$

L'introduction de l'aléa sous cette forme conduit à un processus aléatoire possèdant deux états absorbants : $\{p = 1\}$ et $\{p = 0\}$. Le processus (15) n'est donc plus ergodique et converge, comme chez FH, vers l'un ou l'autre des deux PFS, avec une probabilité égale à un. Dans une telle situation où la stabilité stochastique n'est plus définie, il devient essentiel d'analyser les probabilités de convergence vers les deux PFS. Ces probabilités dépendent des conditions initiales et peuvent conduire à un diagnostic quant à la plausibilité relative des deux équilibres, qui diffère de celui que propose la dynamique déterministe associée. On peut ainsi remarquer dans le cas du modèle BO que la forme probabiliste du processus engendre une dissymètrie entre les deux états absorbants, qui était absente de la dynamique déterministe associée. En effet, l'équation (15) nous dit que :

$$p(t+1) > p(t) \quad \text{ssi} \quad l(t) > g\left(p(t)\right) = \frac{(T-2P)\,p(t) + P}{(R-P)\,p(t)} \quad \text{pour } p(t) \in \left]0, 1\right[\tag{16}$$

Ainsi, si le système se trouve, à l'instant t, près de l'état 1, la probabilité qu'il s'en rapproche à la date suivante, d'après (16), est approximativement égale à $\Pr ob\left(l(t) > g(1) = L_0\right)$. Le même raisonnement appliqué à un voisinage de l'état 0, donne $\Pr ob\left(l(t) < g(\epsilon)\right)$. Comme $g(\epsilon)$ tend vers l'infini quand ϵ tend vers zéro, cette dernière valeur est proche de 1 alors que la précédente peut être notablement inférieure à 1. Cette différence entre les deux états absorbants est d'autant plus forte que L est faible.

5 Conclusion

Les résultats que nous venons de présenter, se rattachent à ce qu'on peut appeler la théorie des systèmes dynamiques soumis à des perturbations aléatoires. Son champ potentiel d'application en économie est immense. Comme l'a montré la notion de stabilité stochastique, elle peut apporter une contribution significative dans la résolution de questions importantes, en l'occurrence, la sélection des équilibres de Nash.

References

[1] BOYER R., ORLEAN A. [1994], ≪ Stabilité de la coopération dans les jeux évolutionnistes stochastiques ≫, *Revue Economique*, 46 (3), p.

797-806.

[2] FOSTER D., YOUNG H.P. [1990], ≪ Stochastic Evolutionary Game Dynamics ≫, *Theoretical Population Biology*, 38, p. 219-232.

[3] FUDENBERG D., HARRIS C. [1992], ≪ Evolutionary Dynamics with Aggregate Shocks ≫, *Journal of Economic Theory*, 57, p. 420-441.

[4] FRIEDMAN D. [1991], ≪ Evolutionary Games in Economics ≫, *Econometrica*, 59 (3), p. 637-666.

[5] KANDORI M., MAILATH G.J., ROB R. [1993], ≪ Learning, Mutation, and Long Run Equilibria in Games ≫, *Econometrica*, 61 (1), p. 29-56.

[6] MAYNARD SMITH J. [1982], *Evolution and the Theory of Games*, Cambridge University Press.