JEUX RÉPÉTÉS À INFORMATION INCOMPLÈTE

par

Jérôme Renault

Table des matières

Dans un jeu répété à information incomplète, les joueurs jouent à chaque étape le même jeu de base, mais celui-ci est imparfaitement connu. On expose ici d'abord le cas le plus simple appelé modèle standard à manque d'information d'un seul côté. Ensuite on présentera quelques variantes et extensions importantes.

Si S est un ensemble fini, on note $|S|$ son cardinal et $\Delta(S)$ l'ensemble des probabilités sur S, identifié à

$$\big\{p = (p_s)_{s\in S} \in \mathbb{R}^S \mid \forall\, s \in S\, p_s \geqslant 0 \text{ et } \textstyle\sum_{s\in S} p_s = 1\big\}.$$

Merci à Dinah Rosenberg pour ses remarques.

Pour $p = (p_s)_{s\in S}$ et $q = (q_s)_{s\in S}$ dans $\mathbb{R}^S$, $\|p-q\|$ désigne, sauf précision contraire, $\sum_{s\in S} |p_s - q_s|$. Pour $C \subset \mathbb{R}^S$, on note int(C) l'intérieur de C, ∂C la frontière de C et conv(C) l'enveloppe convexe de C. Si $G = (G(i,j))_{(i,j)\in I\times J}$ est une matrice de paiement de format $I \times J$, si $x \in \Delta(I)$ et $y \in \Delta(J)$, $G(x,y)$ désigne le paiement espéré

$$\sum_{\substack{i\in I\\ j\in J}} x_i y_j G(i,j).$$

1. Le modèle standard à manque d'information d'un seul côté

Ce modèle a été introduit et étudié par Aumann et Maschler dans les années 1966-68, aux cours de travaux réédités en 1995. On se place dans le cadre des jeux à deux joueurs et à somme nulle, et on fait l'hypothèse que l'un des deux joueurs a toute l'information. Le manque d'information est donc du côté d'un seul joueur. Formellement, on a une famille $(G^k)_{k\in K}$ de jeux matriciels de même taille $I \times J$, et une probabilité p sur K. Le jeu $\Gamma_\infty(p)$ est joué ainsi :

– initialement, un état de la nature k est tiré, une fois pour toutes, selon p. Le joueur 1 apprend k, pas le joueur 2.

– à chaque étape $t = 1, 2, \ldots$, simultanément le joueur 1 choisit une action i_t dans I et le joueur 2 choisit une action j_t dans J. Le paiement d'étape du joueur 1 est alors $G^k(i_t, j_t)$, celui du joueur 2 est $-G^k(i_t, j_t)$, mais tout ce que les joueurs apprennent avant de passer à l'étape $t+1$ est le couple (i_t, j_t).

Le joueur 1 est aussi appelé joueur informé, alors que le joueur 2 est dit non informé. Remarquons que le joueur 1, connaissant k, est capable de déduire son paiement d'étape t ; ce n'est pas le cas du joueur 2 qui n'observe pas son paiement mais uniquement le couple (i_t, j_t). L'ensemble I est appelé l'ensemble d'actions du joueur 1, l'ensemble J celui du joueur 2, et l'ensemble K celui des états de la nature. Ces trois ensembles seront supposés finis et non vides dans tout ce qui suit.

Une stratégie du joueur 1 est un élément $\sigma = (\sigma_t)_{t\geqslant 1}$, avec σ_t : $K \times (I \times J)^{t-1} \to \Delta(I)$ pour tout t (par convention, $(I\times J)^0$ est un singleton). Une stratégie du joueur 2 est un élément $\tau(\tau_t)_{t\geqslant 1}$, avec

$\tau_t : (I \times J)^{t-1} \to \Delta(J)$ pour tout t. On note respectivement Σ et $\mathcal{T}$ les ensembles de stratégies des joueurs 1 et 2. Un couple de stratégies (σ, τ) induit, avec p, une probabilité sur l'ensemble des parties $K \times (I \times J)^{\infty}$ muni de la tribu produit, et pour tout entier strictement positif T on définit le paiement moyen jusqu'à l'étape T par :

$$\gamma_T^p(\sigma, \tau) = \mathbb{E}_{p,\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} G^{\widetilde{k}}(\widetilde{i}_t, \widetilde{j}_t)\right) = \sum_{k \in K} p^k \gamma_T^k(\sigma, \tau).$$

(les « tilde » indiquent des variables aléatoires, et par ailleurs on assimile k et la mesure de Dirac sur k). Le jeu répété T fois est le jeu sous forme stratégique $(\Sigma, \mathcal{T}, \gamma_T^p)$. Par des arguments standards (théorème 2.3 du texte de T. Tomala sur les jeux répétés dans ce volume) il a une valeur notée $v_T(p)$ et les deux joueurs y ont des stratégies optimales. En ce qui concerne la valeur du jeu infiniment répété $\Gamma_\infty(p)$, rappelons la définition suivante :

Définition 1.1. — Soit v un réel.

– Le joueur 1 garantit v dans $\Gamma_\infty(p)$ si :

$$\forall \varepsilon > 0,\ \exists \sigma \in \Sigma,\ \exists T_0,\ \forall T \geqslant T_0,\ \forall \tau \in \mathcal{T}, \quad \gamma_T^p(\sigma, \tau) \geqslant v - \varepsilon.$$

– Le joueur 2 garantit v dans $\Gamma_\infty(p)$ si :

$$\forall \varepsilon > 0,\ \exists \tau \in \mathcal{T},\ \exists T_0,\ \forall T \geqslant T_0,\ \forall \sigma \in \Sigma, \quad \gamma_T^p(\sigma, \tau) \leqslant v + \varepsilon.$$

– v est la valeur du jeu $\Gamma_\infty(p)$ si les deux joueurs y garantissent v.

La valeur du jeu, quand elle existe, est nécessairement unique. Quand le jeu $\Gamma_\infty(p)$ a une valeur v, une stratégie σ du joueur 1 est dite optimale si :

$$\forall \varepsilon > 0,\ \exists T_0,\ \forall T \geqslant T_0,\ \forall \tau \in \mathcal{T}, \quad \gamma_T^p(\sigma, \tau) \geqslant v - \varepsilon,$$

et similairement, une stratégie τ du joueur 2 est dite optimale si :

$$\forall \varepsilon > 0,\ \exists T_0,\ \forall T \geqslant T_0,\ \forall \sigma \in \Sigma, \quad \gamma_T^p(\sigma, \tau) \leqslant v + \varepsilon.$$

1.1. Exemples élémentaires : stratégies complètement révélatrices, non révélatrices et partiellement révélatrices. — Les exemples suivants sont classiques (voir Aumann et Maschler 1995, Zamir 1992). À chaque fois, il y a deux états : $K = \{a, b\}$, et $p = (1/2, 1/2)$.

Exemple 1.2. — $G^a = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$ et $G^b = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$.

Cet exemple est trivial. Afin de maximiser son paiement, le joueur 1 n'a qu'à jouer, à chaque étape, l'action H(aut) si l'état est a et l'action B(as) si l'état est b. Ainsi $v_T(p) = 0 = v_\infty(p) = 0$.

Exemple 1.3. — $G^a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $G^b = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

Une stratégie « naïve » du joueur 1 jouerait à l'étape 1 l'action H si l'état est a, et l'action B si l'état est b. Une telle stratégie est dite *complètement révélatrice*, ou CR, car elle permet au joueur 2 de déduire l'état sélectionné après avoir observé les actions du joueur 1. Cette stratégie est optimale ici dans le jeu à une étape, et $v_1(p) = 1/2$. Mais elle est très mauvaise quand le jeu est répété, et ne garantit rien de plus que 0 dans $\Gamma_\infty(p)$. À l'inverse, le joueur 1 peut toujours ne pas tenir compte de son information, et jouer une stratégie *non révélatrice*, ou NR, *i.e.* jouer indépendamment de l'état. Il considère alors la matrice moyenne

$$\frac{1}{2}G^a + \frac{1}{2}G^b = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix},$$

et peut y jouer à chaque étape une stratégie optimale, qui est ici unique et vaut $\frac{1}{2}H + \frac{1}{2}B$. Ainsi a-t-on : $v_T(p) \geqslant 1/4$ pour tout T. Nous verrons plus tard que cette façon de jouer est ici optimale pour le joueur 1 dans $\Gamma_\infty(p)$.

Exemple 1.4. — $G^a = \begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \end{pmatrix}$ et $G^b = \begin{pmatrix} 0 & 4 & -2 \\ 0 & 4 & 2 \end{pmatrix}$.

Jouer une stratégie CR ne garantit que 0 pour le joueur 1, car le joueur 2 pourra finalement jouer l'action M(du milieu) si l'état est a, et l'action G(auche) si l'état est b. Mais jouer NR revient à se placer dans le jeu

$$\frac{1}{2}G^a + \frac{1}{2}G^b = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \end{pmatrix},$$

et donc ne garantit que 0. Nous prouverons plus tard qu'il est ici optimal pour le joueur 1 de jouer la stratégie σ présentée maintenant.

Le joueur 1 choisit aléatoirement, une fois pour toutes, un élément s dans $\{H, B\}$ de la façon suivante : si $k = a$, alors $s = H$ avec probabilité $3/4$, et donc $s = B$ avec probabilité $1/4$; et si $k = b$, alors

$s = H$ avec probabilité $1/4$, et $s = B$ avec probabilité $3/4$. Ensuite le joueur 1 joue l'action s à chaque étape, indépendamment des coups du joueur 2.

Les probabilités conditionnelles vérifient : $P(k = a|s = H) = 3/4$, et $P(k = a|s = B) = 1/4$. Donc à la fin de l'étape 1, le joueur 2 ayant observé le premier coup du joueur 1 aura appris quelque chose sur l'état de la nature : sa croyance sera passée de $\frac{1}{2}a + \frac{1}{2}b$ à $\frac{3}{4}a + \frac{1}{4}b$ ou à $\frac{1}{4}a + \frac{3}{4}b$. Mais il ne connaît toujours pas l'état avec probabilité 1 : on parle de *révélation partielle d'information.*

1.2. Utilisation de l'information : éclatement et concavification

$p = (p^k)_{k\in K}$ est la probabilité initiale, ou l'*a priori* du joueur 2 sur l'état de la nature. Le joueur 1 choisit sa première action (ou plus généralement un message ou signal s au sein d'un ensemble fini S), en fonction de l'état de la nature k sélectionné. Notons $x = (x^k)_{k\in K} \in \Delta(S)^K$ la probabilité de transition utilisée : si l'état est k, le joueur 1 choisit le signal s avec probabilité $x^k(s)$.

La probabilité totale que s arrive est notée : $\lambda_s = \sum_{k\in K} p^k x^k(s)$, et si $\lambda_s > 0$, la probabilité conditionnelle sur K (ou *a posteriori*) sachant s vaut : $\widehat{p}(x,s) = \left(\frac{p^k x^k(s)}{\lambda_s}\right)_{k\in K}$. Il est clair que :

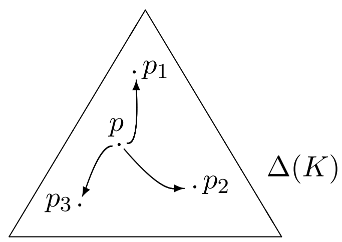
$$\sum_{s\in S} \lambda_s \widehat{p}(x,s) = p. \tag{1}$$

Les *a posteriori* contiennent donc l'*a priori* p dans leur enveloppe convexe. Le lemme suivant est fondamental et exprime une sorte de réciproque : on dit que le joueur 1 peut amener n'importe quels *a posteriori* contenant la probabilité initiale dans leur enveloppe convexe.

***Lemme 1.5* (éclatement).** — *Supposons que $p = \sum_{s\in S} \lambda_s p_s$, avec S fini, pour tout s de S $\lambda_s > 0$, $p_s \in \Delta(K)$, et $\sum_{s\in S} \lambda_s = 1$. Alors il existe une probabilité de transition $x \in \Delta(S)^K$ telle que :*

$$\forall\, s \in S, \quad \lambda_s = \sum_{k\in K} p^k x^k(s) \quad \textit{et} \quad \widehat{p}(x,s) = p_s.$$

Démonstration. — On pose $x^k(s) = \lambda_s p_s^k / p^k$ si $p^k > 0$. □



Soit f une application semi-continue supérieurement de $\Delta(K)$ dans $\mathbb{R}$. La plus petite fonction concave de $\Delta(K)$ dans $\mathbb{R}$ partout plus grande que f, notée $\operatorname{cav} f$, est continue et, pour tout p dans $\Delta(K)$,

$$\begin{aligned}\operatorname{cav} f(p) = \max \Big\{ & \textstyle\sum_{s \in S} \lambda_s f(p_s) \mid S \text{ fini}, \forall\, s \in S\; \lambda_s \geqslant 0, \\ & p_s \in \Delta(K), \textstyle\sum_{s \in S} \lambda_s = 1, \textstyle\sum_{s \in S} \lambda_s p_s = p \Big\}.\end{aligned}$$

Nous allons voir que le lemme 1.5 a une conséquence importante du point de vue de ce qui peut être garanti par le joueur 1. Le résultat suivant est valable dans n'importe quel jeu finiment ou infiniment répété (en fait dans tout jeu à somme nulle où le joueur 1 a toute l'information). On dit que le joueur 1 garantit f si pour chaque valeur de p, le joueur 1 garantit $f(p)$ dans le jeu de probabilité initiale p.

Lemme 1.6. — *Si le joueur 1 garantit f, alors le joueur 1 garantit* $\operatorname{cav} f$.

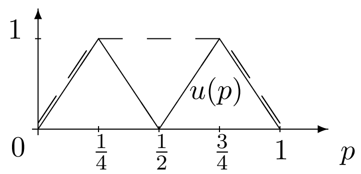
Démonstration. — Fixons p, et considérons l'égalité $\operatorname{cav} f(p) = \sum_{s \in S} \lambda_s f(p_s)$, avec les notations précédentes. Le joueur 1 observe l'état k, puis choisit s dans S selon $x^k(s)$, où $x = (x^k)_{k \in K}$ est donné par le splitting (lemme 1.5). Puis il joue de façon à garantir $f(p_s)$ dans le jeu de probabilité initiale p_s. Par le théorème de Kuhn sur l'équivalence entre les stratégies générales et les stratégies de comportement (voir le théorème 1.4 du texte de T. Tomala sur les jeux répétés dans ce volume), la stratégie du joueur 1 ainsi définie peut se voir comme une stratégie dans Σ. Ainsi le joueur 1 garantit-il $\sum_{s \in S} \lambda_s f(p_s)$ dans le jeu de probabilité initiale p. □

Remarquons que le fait que le joueur 2 observe ou non, apprenne ou pas, l'élément s n'a aucune influence sur la preuve ci-dessus. Dans un jeu à somme nulle où le joueur 1 (maximisateur) a toute l'information, le lemme 1.6 donne : si le jeu a une valeur $v(p)$ pour toute probabilité initiale p, alors v est nécessairement concave. On peut voir cela comme le fait que la valeur de l'information est positive dans un jeu à somme nulle.

1.3. Le joueur 1 garantit $\operatorname{cav} u$**.** — Notons, pour toute probabilité p dans $\Delta(K)$, $u(p)$ la valeur du jeu matriciel $G(p) = \sum_{k\in K} p^k G^k$. Le joueur 1 peut garantir $u(p)$ dans $\Gamma_\infty(p)$ en jouant à chaque étape une stratégie optimale dans $G(p)$, indépendamment de l'état et des coups passés. u est la valeur du jeu répété où le joueur 1 joue indépendamment de l'état, et $G(p)$ s'appelle le jeu non révélateur à p. Clairement le joueur 1 garantit u, donc on a, dans un jeu à T étapes ou dans le jeu infiniment répété :

Corollaire 1.7. — *Le joueur 1 garantit* $\operatorname{cav} u$.

Revenons sur les exemples précédents. Dans l'exemple 1.2, on a $u(p) = -p(1-p)$ pour tout p (en faisant l'abus de notation : p probabilité de l'état a). u est convexe, et $\operatorname{cav} u(p) = 0$ pour tout p. Dans l'exemple 1.3, on a $u(p) = p(1-p)$ pour tout p, donc u est concave et $u = \operatorname{cav} u$. Concernant l'exemple 1.4, on a représenté u, en plein, et $\operatorname{cav} u$, en pointillés, sur la figure suivante.



Reprenons la stratégie partiellement révélatrice du joueur 1 présentée précédemment. Avec probabilité $1/2$, l'*a posteriori* vaudra $\frac{3}{4}a + \frac{1}{4}b$, et le joueur 1 jouera H qui est optimale dans

$$\frac{3}{4}G^a + \frac{1}{4}G^b = \begin{pmatrix} 3 & 1 & 1 \\ 3 & 1 & -1 \end{pmatrix}.$$

De même avec probabilité $1/2$ l'*a posteriori* vaudra $\frac{1}{4}a + \frac{3}{4}b$ et le joueur 1 jouera une stratégie optimale dans $\frac{1}{4}G^a + \frac{3}{4}G^b$. Cette stratégie garantit donc :

$$\frac{1}{2}\,u(3/4) + \frac{1}{2}\,u(1/4) = \operatorname{cav} u(1/2).$$

1.4. Le joueur 2 garantit $\lim_T v_T$**.** — T étant fixé, le jeu finiment répété à T étapes $\Gamma_T(p)$ a une valeur $v_T(p)$ et les deux joueurs y ont des stratégies optimales. v_T est concave d'après le lemme 1.6, et lipschitzienne de constante

$$M = \max\{|G^k(i,j)| \mid k \in K,\, i \in I,\, j \in J\}$$

indépendante de T.

Définition 1.8. — Pour p dans $\Delta(K)$, on pose $v^*(p) = \inf_{T\geqslant 1} v_T(p)$.

v^* est concave et M-Lipschitz.

Lemme 1.9. — *La suite de fonctions $(v_T)_T$ converge uniformément sur $\Delta(K)$ vers v^*, et le joueur 2 garantit v^* dans le jeu infiniment répété.*

Démonstration. — Le joueur 2 peut se comporter ainsi dans le jeu $\Gamma_\infty(p)$: jouer une stratégie optimale de $\Gamma_T(p)$, puis tout oublier, et recommencer à jouer une stratégie optimale de $\Gamma_T(p)$, puis tout oublier, et recommencer etc. Ainsi le joueur 2 garantit $v_T(p)$ dans $\Gamma_\infty(p)$. Donc le joueur 2 garantit $v^*(p)$. D'après la définition 1.1, cela implique que : $\limsup_T v_T(p) \leqslant v^*(p)$. Ce qui entraîne que v_T converge simplement vers v^*. La convergence est uniforme car toutes ces fonctions sont lipschitziennes de même constante. □

Indiquons également que l'on a pour tous T, $T' \geqslant 1$, et p dans $\Delta(K)$:

$$(T+T')\, v_{T+T'}(p) \leqslant T\, v_T(p) + T'\, v_{T'}(p).$$

En effet, le joueur 2 peut jouer une stratégie optimale dans $\Gamma_T(p)$, puis tout oublier et jouer une stratégie optimale dans $\Gamma_{T'}(p)$.

1.5. Martingale des *a posteriori*. — La définition suivante est essentielle.

Définition 1.10. — Soit (σ,τ) dans $\Sigma\times\mathcal{T}$ un profil de stratégies dans $\Gamma_\infty(p)$. Pour t dans $\mathbb{N}$ et $h_t = (i_1, j_1, \ldots, i_t, j_t) \in (I\times J)^t$, on définit l'*a posteriori* du joueur 2 après h_t comme :

$$p_t(\sigma,\tau,h_t) = (p_t^k(\sigma,\tau,h_t))_{k\in K} \in \Delta(K),$$

avec $p_t^k(\sigma,\tau,h_t) = \mathbb{P}_{p,\sigma,\tau}(\widetilde{k} = k | h_t)$ pour tout k.

$p_t(\sigma, \tau, h_t)$ est la croyance du joueur 2 sur l'état de la nature à la fin de l'étape t si h_t a été précédemment joué et que le joueur 1 utilise σ. On a facilement que $p_t(\sigma, \tau, h_t)$ ne dépend ni de τ, ni de l'action jouée par le joueur 2 en date t, pourvu que $\mathbb{P}_{p,\sigma,\tau}(h_t) > 0$. On note donc aussi $p_t(\sigma, h_t)$ pour $p_t(\sigma, \tau, h_t)$, dès qu'il existe τ telle que $\mathbb{P}_{p,\sigma,\tau}(h_t) > 0$ (et on peut définir arbitrairement $p_t(\sigma, \tau, h_t)$ et $p_t(\sigma, h_t)$ dans $\Delta(K)$ sinon). Notons $H_t = (I \times J)^t$ et $\mathcal{H}_t$ la tribu sur $K \times (I \times J)^\infty$ engendrée par la projection sur H_t donnant les actions jouées aux t premières étapes. $p_t(\sigma)$ est vue comme une application mesurable de $(K \times (I \times J)^\infty, \mathcal{H}_t)$ dans $\Delta(K)$, donc est une variable aléatoire à valeurs dans $\Delta(K)$. $p_0(\sigma)$ est la probabilité initiale p. La propriété suivante est cruciale.

Propriété 1.11. — Par rapport à $\mathbb{P}_{p,\sigma,\tau}$, la suite $(p_t(\sigma))_{t\geqslant 0}$ est une $(\mathcal{H}_t)_{t\geqslant 0}$- martingale à valeurs dans $\Delta(K)$.

Cette propriété est très générale et concerne l'apprentissage bayésien d'un paramètre inconnu : l'espérance de ce que je saurai demain est ce que je sais aujourd'hui. On peut facilement en donner une démonstration analytique de façon analogue à l'obtention de l'équation (1).

Fixons maintenant une stratégie σ du joueur 1. L'idée des calculs suivants est la suivante. La martingale $(p_t(\sigma))_{t\geqslant 0}$ étant bornée, elle converge p.s. et on a une borne uniforme sur sa variation L^1 (voir le lemme 1.12). Cela implique qu'au bout d'un moment, la martingale sera essentiellement constante, et donc qu'à partir d'une certaine étape le joueur 1 jouera approximativement de façon non révélatrice, et ne pourra pas garantir beaucoup plus que $u(q)$, où q est « un *a posteriori* limite ». Les *a posteriori* étant liés à la probabilité initiale, le joueur 1 ne pourra pas garantir plus, sur le long terme, que

$$\max\Big\{ \textstyle\sum_{s\in S} \lambda_s u(p_s) \mid S \text{ fini}, \forall\, s \in S\, \lambda_s \geqslant 0, \\ p_s \in \Delta(K), \textstyle\sum_{s\in S} \lambda_s = 1, \textstyle\sum_{s\in S} \lambda_s p_s = p\Big\},$$

autrement dit que $\operatorname{cav} u(p)$. Passons maintenant à la preuve formelle.

σ étant fixée, on définit la stratégie $\tau = (\tau_t)_{t\geqslant 1}$ du joueur 2 de la façon suivante : jouer à chaque étape une stratégie optimale dans le jeu matriciel $G(p_t)$, où p_t est l'a posteriori courant.

Supposons que (σ,τ) soit joué dans le jeu répété $\Gamma_\infty(p)$. Afin d'alléger les notations, on note dans les calculs suivants $\mathbb{P}=\mathbb{P}_{p,\sigma,\tau}$, on note $\mathbb{E}$ l'espérance par rapport à $\mathbb{P}$, et $p_t(h_t)$ pour $p_t(\sigma,h_t)$. Toutes les normes indiquées sont des normes 1. L'inégalité suivante est juste due au fait que $(p_t)_t$ est une martingale à valeurs dans $\Delta(K)$ et d'espérance p.

Lemme 1.12

$$\forall\, T\geqslant 1,\quad \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}(\|p_{t+1}-p_t\|)\leqslant\frac{\sum_{k\in K}\sqrt{p^k(1-p^k)}}{\sqrt{T}}.$$

Démonstration. — On a pour tous k dans K et $t\geqslant 0$:

$$\mathbb{E}\big((p^k_{t+1}-p^k_t)^2\big)=\mathbb{E}\Big(\mathbb{E}\big((p^k_{t+1}-p^k_t)^2|\mathcal{H}_t\big)\Big)=\mathbb{E}\big((p^k_{t+1})^2\big)-\mathbb{E}\big((p^k_t)^2\big).$$

Donc

$$\mathbb{E}\bigg(\sum_{t=0}^{T-1}(p^k_{t+1}-p^k_t)^2\bigg)=\mathbb{E}\big((p^k_T)^2\big)-(p^k)^2\leqslant p^k(1-p^k).$$

Comme par Cauchy-Schwartz

$$\mathbb{E}\bigg(\frac{1}{T}\sum_{t=0}^{T-1}\Big|p^k_{t+1}-p^k_t\Big|\bigg)\leqslant\sqrt{\frac{1}{T}\mathbb{E}\bigg(\sum_{t=0}^{T-1}(p^k_{t+1}-p^k_t)^2\bigg)}$$

pour chaque k, on a le résultat voulu. $\square$

Pour h_t dans H_t, on note $\sigma^k(h_t)=\sigma_{t+1}(k,h_t)\in\Delta(I)$ l'action mixte jouée par le joueur 1 après h_t si l'état est k, et $\overline{\sigma}(h_t)$ la loi de l'action de date $t+1$ jouée par le joueur 1 après h_t :

$$\overline{\sigma}(h_t)=\sum_{k\in K}p^k_t(\sigma,h_t)\sigma^k(h_t)\in\Delta(I).$$

$\overline{\sigma}(h_t)$ peut se voir comme la stratégie moyenne après h_t, et servira d'approximation non révélatrice à $(\sigma^k(h_t))_k$. Le lemme suivant exprime le lien entre, d'une part, la variation de la martingale $(p_t(\sigma))_{t\geqslant 0}$, *i.e.* l'information révélée par le joueur 1, et d'autre part, la dépendance par rapport à l'état de l'action jouée par le joueur 1 en date $t+1$, *i.e.* l'information employée par le joueur 1.

Lemme 1.13. — *$\forall\, t\geqslant 0$, $\forall\, h_t\in H_t$,*

$$\mathbb{E}\left(\|p_{t+1}-p_t\|\,|h_t\right)=\mathbb{E}\Big(\big\|\sigma^{\widetilde{k}}(h_t)-\overline{\sigma}(h_t)\big\|\big|h_t\Big).$$

Démonstration. — Fixons $t \geqslant 0$ et h_t dans H_t tel que $\mathbb{P}_{p,\sigma,\tau}(h_t) > 0$. Pour (i_{t+1}, j_{t+1}) dans $I \times J$, on a :

$$\begin{aligned} p^k_{t+1}(h_t, i_{t+1}, j_{t+1}) &= \mathbb{P}(\widetilde{k} = k | h_t, i_{t+1}) \\ &= \frac{\mathbb{P}(\widetilde{k} = k | h_t)\mathbb{P}(i_{t+1} | k, h_t)}{\mathbb{P}(i_{t+1} | h_t)} \\ &= \frac{p^k_t(h_t)\sigma^k(h_t)(i_{t+1})}{\overline{\sigma}(h_t)(i_{t+1})}. \end{aligned}$$

Donc :

$$\begin{aligned} \mathbb{E}(\|p_{t+1} - p_t\| | h_t) &= \sum_{i_{t+1}\in I} \overline{\sigma}(h_t)(i_{t+1}) \sum_{k\in K} |p^k_{t+1}(h_t, i_{t+1}) - p^k_t(h_t)| \\ &= \sum_{i_{t+1}\in I} \sum_{k\in K} |p^k_t(h_t)\sigma^k(h_t)(i_{t+1}) - \overline{\sigma}(h_t)(i_{t+1})p^k_t(h_t)| \\ &= \sum_{k\in K} p^k_t(h_t)\|\sigma^k(h_t) - \overline{\sigma}(h_t)\| \\ &= \mathbb{E}\big(\|\sigma^{\widetilde{k}}(h_t) - \overline{\sigma}(h_t)\| | h_t\big). \end{aligned}$$

□

On peut maintenant majorer les paiements. Pour $t \geqslant 0$ et h_t dans H_t :

$$\begin{aligned} \mathbb{E}\left(G^{\widetilde{k}}(\widetilde{i}_{t+1}, \widetilde{j}_{t+1}) | h_t\right) &= \sum_{k\in K} p^k_t(h_t)G^k(\sigma^k(h_t), \tau_{t+1}(h_t)) \\ &\leqslant \sum_{k\in K} p^k_t(h_t)G^k(\overline{\sigma}(h_t), \tau_{t+1}(h_t)) \\ &\qquad + M \sum_{k\in K} p^k_t(h_t)\|\sigma^k(h_t) - \overline{\sigma}(h_t)\| \\ &\leqslant u(p_t(h_t)) + M \sum_{k\in K} p^k_t(h_t)\|\sigma^k(h_t) - \overline{\sigma}(h_t)\|, \end{aligned}$$

où $u(p_t(h_t))$ vient de la définition de τ. Par le lemme 1.13, on obtient :

$$\mathbb{E}\left(G^{\widetilde{k}}(\widetilde{i}_{t+1}, \widetilde{j}_{t+1}) | h_t\right) \leqslant u(p_t(h_t)) + M\mathbb{E}\left(\|p_{t+1} - p_t\| | h_t\right).$$

En appliquant l'inégalité de Jensen, il vient :

$$\mathbb{E}\left(G^{\widetilde{k}}(\widetilde{i}_{t+1}, \widetilde{j}_{t+1})\right) \leqslant \operatorname{cav} u(p) + M\mathbb{E}\left(\|p_{t+1} - p_t\|\right).$$

On applique maintenant le lemme 1.12, et on obtient :

$$\begin{aligned}\gamma_T^p(\sigma,\tau) &= \mathbb{E}\bigg(\frac{1}{T}\sum_{t=0}^{T-1} G^{\widetilde{k}}(\widetilde{i}_{t+1},\widetilde{j}_{t+1})\bigg)\\ &\leqslant \operatorname{cav} u(p) + \frac{M}{\sqrt{T}}\sum_{k\in K}\sqrt{p^k(1-p^k)}.\end{aligned}$$

Il reste à considérer le cas d'une stratégie σ optimale dans $\Gamma_T(p)$ et on a prouvé le résultat suivant.

Proposition 1.14. — *Pour p dans $\Delta(K)$ et $T \geqslant 1$,*

$$v_T(p) \leqslant \operatorname{cav} u(p) + \frac{M'}{\sqrt{T}}, \quad \textit{avec} \quad M' = M\sum_{k\in K}\sqrt{p^k(1-p^k)}.$$

1.6. Valeur de $\Gamma_\infty(p)$. — D'après le corollaire 1.7, pour tout T le joueur 1 garantit $\operatorname{cav} u(p)$ dans $\Gamma_T(p)$, donc on a $v_T(p) \geqslant \operatorname{cav} u(p)$, et donc $v^* \geqslant \operatorname{cav} u$. En passant à la limite en T dans la proposition 1.14, on trouve $v^* \leqslant \operatorname{cav} u$, donc $v^* = \operatorname{cav} u$. Avec le corollaire 1.7 et le lemme 1.9, on obtient :

***Théorème 1.15* (Aumann et Maschler).** — *Le jeu $\Gamma_\infty(p)$ a une valeur qui vaut* $\operatorname{cav} u(p)$.

1.7. Formule de récurrence. — On montre facilement la formule de récurrence suivante, où $\max_{x\in\Delta(I)^K}$ et $\min_{y\in\Delta(J)}$ commutent par le théorème de Sion.

$$\begin{aligned}&v_{T+1}(p)\\ &\quad= \frac{1}{T+1}\max_{x\in\Delta(I)^K}\min_{y\in\Delta(J)}\bigg(G(p,x,y) + T\sum_{i\in I} x(p)(i) v_T(\widehat{p}(x,i))\bigg),\end{aligned}$$

où, avec des notations similaires au paragraphe 1.2 :

– $x = (x^k(i))_{i\in I, k\in K}$, avec x^k la probabilité utilisée par le joueur 1 à l'étape 1 si l'état est k,

– $y = (y(j))_{j\in J}$ est la probabilité utilisée par le joueur 2 à l'étape 1,

– $G(p,x,y) = \sum_k p^k G^k(x^k, y)$ est le paiement espéré d'étape 1,

– $x(p)(i) = \sum_{k\in K} p^k x^k(i)$ est la probabilité que i soit jouée en date 1,

– $\widehat{p}(x,i)$ est la probabilité conditionnelle sur K sachant i.

La propriété suivante s'interprète facilement : plus on joue longtemps, moins l'information initiale n'est importante.

Lemme 1.16. — *$(v_T(p))_T$ est décroissante.*

Démonstration. — Montrons $v_{T+1} \leqslant v_T$ par récurrence sur T.

$$v_2(p) = \frac{1}{2} \max_{x\in\Delta(I)^K} \min_{y\in\Delta(J)} \Big(G(p,x,y) + T \sum_{i\in I} x(p)(i) v_1(\widehat{p}(x,i)) \Big)$$

Comme v_1 est concave,

$$v_2(p) \leqslant \frac{1}{2} \max_{x\in\Delta(I)^K} \min_{y\in\Delta(J)} (G(p,x,y) + v_1(p)) = v_1(p).$$

Soit $T \geqslant 2$ tel que : $\forall\, p$, $v_T(p) \leqslant v_{T-1}(p)$. Alors,

$$\begin{aligned}
(T+1)v_{T+1}(p) = {} & \max_{x\in\Delta(I)^K} \min_{y\in\Delta(J)} \Big(G(p,x,y) \\
& + (T-1) \sum_{i\in I} x(p)(i) v_T(\widehat{p}(x,i)) + \sum_{i\in I} x(p)(i) v_T(\widehat{p}(x,i) \Big) \\
& \leqslant \max_{x\in\Delta(I)^K} \min_{y\in\Delta(J)} \Big(G(p,x,y) \\
& \qquad + (T-1) \sum_{i\in I} x(p)(i) v_{T-1}(\widehat{p}(x,i)) + v_T(p) \Big) \\
& = T v_T(p) + v_T(p) \\
& = (T+1) v_T(p).
\end{aligned}$$

□

2. Jeux à paiements vectoriels et approchabilité

Le modèle suivant a été introduit par D. Blackwell (1956). Il est intéressant en tant que tel, mais permettra aussi d'expliciter, en 2.4, la construction d'une stratégie optimale pour le joueur non informé dans le jeu $\Gamma_\infty(p)$ de la section précédente. On a ici aussi une famille de matrices $(G^k)_{k\in K}$ de même taille $I \times J$. À chaque étape t, le joueur 1 choisit $i_t \in I$, simultanément le joueur 2 choisit $j_t \in J$, et le « paiement » d'étape t est alors le vecteur $G(i_t, j_t) = (G^k(i_t, j_t))_{k\in K}$ dans $\mathbb{R}^K$. Précisons qu'il n'y a pas de vrai état k ici, ni de probabilité a priori sur K, et les deux joueurs ont un rôle symétrique. On suppose qu'après chaque étape chaque joueur observe le vecteur de paiement. Les actions des joueurs ne sont en principe pas observées (on pourra remarquer plus tard que cela ne changerait en fait rien aux résultats, qui nécessitent juste de supposer que les joueurs observent *au moins* le vecteur de paiement). L'approchabilité vise à répondre aux questions du genre : dans quels ensembles le joueur 1 (par exemple) peut-il amener le paiement moyen de long terme ?

Dans toute cette section sur l'approchabilité, on considère norme et distance euclidiennes. La présentation s'est inspirée de Sorin (2002). Notons $F = \{(G^k(i,j))_{k\in K} \mid i \in I,\, j \in J\}$ l'ensemble fini des paiements d'étapes possibles, et M une constante telle que $\|u\| \leqslant M$ pour tout u de F. Une stratégie du joueur 1 est un élément $\sigma = (\sigma_t)_{t\geqslant 1}$, où σ_t est une application de F^{t-1} dans $\Delta(I)$ pour tout t. De même pour le joueur 2 en remplaçant $\Delta(I)$ par $\Delta(J)$. On note respectivement Σ et $\mathcal{T}$ les espaces de stratégies des joueurs 1 et 2. Un profil de stratégies (σ,τ) dans $\Sigma\times\mathcal{T}$ induit « naturellement » une unique probabilité sur $(I\times J\times F)^{\infty}$ notée $\mathbb{P}_{\sigma,\tau}$. On note g_t la variable aléatoire, à valeurs dans F, du paiement d'étape t, et $\overline{g}_t = \frac{1}{t}\sum_{t'=1}^{t} g_{t'} \in \mathrm{conv}(F)$. Soit $C \subset \mathbb{R}^K$ un ensemble « cible », que l'on supposera toujours sans perte de généralité fermé. On note $d_t = d(\overline{g}_t, C)$ la v.a. de la distance euclidienne de $\overline{g}_t$ à C.

Définition 2.1

– C est approchable par le joueur 1 si :

$$\forall\,\varepsilon>0,\ \exists\,\sigma\in\Sigma,\ \exists\,T,\ \forall\,\tau\in\mathcal{T},\ \forall\,t\geqslant T,\quad \mathbb{E}_{\sigma,\tau}(d_t)\leqslant\varepsilon.$$

– C est approchable par le joueur 2 si :

$$\forall\,\varepsilon>0,\ \exists\,\tau\in\mathcal{T},\ \exists\,T,\ \forall\,\sigma\in\Sigma,\ \forall\,t\geqslant T,\quad \mathbb{E}_{\sigma,\tau}(d_t)\leqslant\varepsilon.$$

– C est évitable (on dit aussi repoussable) par le joueur 1, resp. joueur 2, s'il existe $\delta>0$ tel que $\{z\in\mathbb{R}^K, d(z,C)\geqslant\delta\}$ soit approchable par le joueur 1, resp. joueur 2.

C est approchable par le joueur 1 si pour tout $\varepsilon>0$, le joueur 1 peut s'assurer qu'au bout d'un certain temps on aura $\mathbb{E}(d_t)\leqslant\varepsilon$, donc on sera en espérance proche de C à ε près. Un ensemble ne peut être à la fois approchable par le joueur 1 et évitable par le joueur 2. Par exemple si K est un singleton, on est en dimension 1 et le théorème du minmax implique que pour tout réel t, l'ensemble $[t,+\infty[$ est soit approchable par le joueur 1, soit évitable par le joueur 2 (selon la position de t par rapport à $\max_{x\in\Delta(I)}\min_{y\in\Delta(J)} G(x,y)$).

2.1. Conditions suffisantes et nécessaires d'approchabilité

Pour x dans $\Delta(I)$, on note

$$xG = \{G(x, y) \mid y \in \Delta(J)\} = \operatorname{conv}\{\textstyle\sum_{i \in I} x_i G(i, j) \mid j \in J\}.$$

xG est l'ensemble des paiements vectoriels espérés possibles quand le joueur 1 joue l'action mixte x. De même, pour y dans $\Delta(J)$ on note

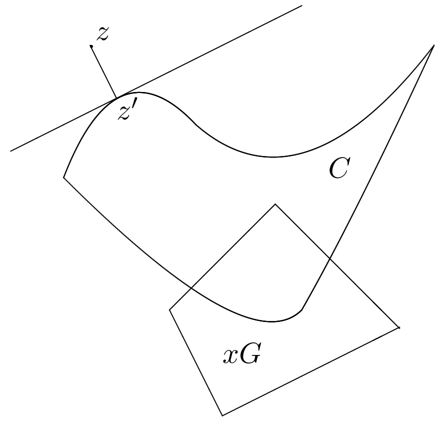
$$Gy = \{G(x, y) \mid x \in \Delta(I)\}.$$

Dans la définition suivante, la lettre B fait référence à Blackwell.

Définition 2.2. — C est un B-ensemble pour le joueur 1 si pour tout $z \notin C$, il existe $z' \in C$ et $x \in \Delta(I)$ satisfaisant les deux conditions suivantes :

(i) $\|z' - z\| = d(z, C)$,

(ii) L'hyperplan affine passant par z' et orthogonal à $[z, z']$ sépare z de xG.



Par exemple, si x est dans $\Delta(I)$, l'ensemble convexe xG est approchable par le joueur 1. Étant donné un B-ensemble C pour le joueur 1, on définit une stratégie σ adaptée à C de la façon suivante. Pour tout t de $\mathbb{N}$, en date $t+1$ le joueur 1 considère le paiement moyen courant $\overline{g}_t$. Si $\overline{g}_t \in C$ (ou si $t = 0$), σ joue arbitrairement en date $t + 1$. Si $\overline{g}_t \notin C$ (et $t \geqslant 1$), σ joue en date $t + 1$ une action mixte $x \in \Delta(I)$ vérifiant la définition précédente pour $z = \overline{g}_t$.

Théorème 2.3. — *Soit C un B-ensemble pour le joueur 1. Une stratégie σ adaptée à C vérifie :*

$$\forall\,\tau \in \mathcal{T},\, \forall\, t \geqslant 1 \quad \mathbb{E}_{\sigma,\tau}(d_t) \leqslant \frac{2M}{\sqrt{t}} \quad \textit{et} \quad d_t \underset{t\to\infty}{\longrightarrow} 0 \quad \mathbb{P}_{\sigma,\tau}\, p.s.$$

En dimension 1 et pour $C = \{0\}$, ce théorème implique en particulier qu'une suite de réels $(x_t)_t$ bornée, telle que le produit

$$x_{T+1}\biggl(\frac{1}{T}\sum_{t=1}^{T} x_T\biggr)$$

soit négatif pour tout T, converge en moyenne de Césaro vers 0.

Démonstration. — Supposons que le joueur 1 joue σ adaptée à C, alors que le joueur 2 joue une stratégie quelconque τ. Soit $t \geqslant 1$, et supposons que $\overline{g}_t \notin C$. On note $z' \in C$ et $x \in \Delta(I)$ qui satisfont (i) et (ii) de la définition 2.2 pour $z = \overline{g}_t$. On a :

$$\begin{aligned} d_{t+1}^2 = d(\overline{g}_{t+1}, C)^2 \leqslant \|\overline{g}_{t+1} - z'\|^2 & \\ &= \Bigl\|\frac{1}{t+1}\sum_{l=1}^{t+1} g_l - z'\Bigr\|^2 \\ &= \Bigl\|\frac{1}{t+1}(g_{t+1} - z') + \frac{t}{t+1}(\overline{g}_t - z')\Bigr\|^2 \\ &= \Bigl(\frac{1}{t+1}\Bigr)^2 \|g_{t+1} - z'\|^2 + \Bigl(\frac{t}{t+1}\Bigr)^2 {d_t}^2 \\ &\qquad + \frac{2t}{(t+1)^2}\langle g_{t+1} - z', \overline{g}_t - z'\rangle. \end{aligned}$$

Par hypothèse, l'espérance, sachant les t premiers coups $h_t \in (I \times J)^t$, du produit scalaire ci-dessus est négatif, donc

$$\mathbb{E}\left(d_{t+1}^2 | h_t\right) \leqslant \Bigl(\frac{t}{t+1}\Bigr)^2 {d_t}^2 + \Bigl(\frac{1}{t+1}\Bigr)^2 \mathbb{E}\left(\|g_{t+1} - z'\|^2 | h_t\right).$$

Or

$$\mathbb{E}\left(\|g_{t+1} - z'\|^2 | h_t\right) \leqslant \mathbb{E}\Bigl(\|g_{t+1} - \overline{g}_t\|^2 | h_t\Bigr) \leqslant (2M)^2,$$

donc :

$$\mathbb{E}\left(d_{t+1}^2 | h_t\right) \leqslant \Bigl(\frac{t}{t+1}\Bigr)^2 {d_t}^2 + \Bigl(\frac{1}{t+1}\Bigr)^2 4M^2. \tag{2}$$

En prenant l'espérance, on obtient, que $\overline{g}_t \notin C$ ou pas :

$$\forall\, t \geqslant 1, \quad \mathbb{E}\left(d_{t+1}^2\right) \leqslant \Bigl(\frac{t}{t+1}\Bigr)^2 \mathbb{E}({d_t}^2) + \Bigl(\frac{1}{t+1}\Bigr)^2 4M^2.$$

Donc on a par récurrence, pour tout $t \geqslant 1$, $\mathbb{E}(d_t^2) \leqslant 4M^2/t$, puis

$$\mathbb{E}(d_t) \leqslant \frac{2M}{\sqrt{t}}.$$

Posons maintenant, comme dans Sorin (2002),

$$e_t = d_t^2 + \sum_{t'>t} \frac{4M^2}{t'^2}.$$

L'inégalité (2) donne :

$$\mathbb{E}(e_{t+1}|h_t) \leqslant e_t,$$

donc (e_t) est une surmartingale positive dont l'espérance tend vers zéro. Donc $e_t \to_{t\to\infty} 0$ $\mathbb{P}_{\sigma,\tau}$ p.s., et enfin $d_t \to_{t\to\infty} 0$ $\mathbb{P}_{\sigma,\tau}$ p.s. □

Le théorème 2.3 implique que tout B-ensemble pour le joueur 1 est approchable par le joueur 1. La réciproque est vraie dans le cas des ensembles convexes.

Théorème 2.4. — *Soit C un convexe fermé de $\mathbb{R}^K$.*

$$\begin{aligned} & \text{(i) } C \textit{ est un } B\textit{-ensemble pour le joueur 1,} \\ \iff & \text{(ii) } \forall\, y \in \Delta(J), Gy \cap C \neq \varnothing, \\ \iff & \text{(iii) } C \textit{ est approchable par le joueur 1,} \\ \iff & \text{(iv) } \forall\, q \in \mathbb{R}^K,\ \max_{x\in\Delta(I)} \min_{y\in\Delta(J)} \sum_{k\in K} q^k G^k(x,y) \geqslant \inf_{c\in C} \langle q, c\rangle. \end{aligned}$$

Démonstration. — (i) ⇒ (iii) vient du théorème précédent.

Montrons (iii) ⇒ (ii). Supposons qu'il existe $y \in \Delta(J)$ tel que $Gy \cap C = \varnothing$. Comme Gy est approchable par le joueur 2, alors C est évitable par le joueur 2 et donc pas approchable par le joueur 1.

Montrons (ii) ⇒ (i). Supposons que $Gy \cap C \neq \varnothing$ pour tout $y \in \Delta(J)$. Soient $z \notin C$ et z' sa projection sur C. Considérons le jeu matriciel où les paiements de G sont projetés dans la direction $z' - z$, *i.e.* le jeu matriciel $\sum_{k\in K}(z'^k - z^k)G^k$. Par hypothèse, on a :

$$\forall\, y \in \Delta(J), \exists\, x \in \Delta(I), \quad G(x,y) \in C,$$

donc :

$$\langle z' - z, G(x,y)\rangle \geqslant \min_{c\in C}\{\langle z' - z, c\rangle, \langle z' - z, z'\rangle\}.$$

Donc

$$\min_{y\in\Delta(J)} \max_{x\in\Delta(I)} \langle z' - z, G(x,y)\rangle \geqslant \langle z' - z, z'\rangle.$$

Par le théorème du minmax, il existe x dans $\Delta(I)$ tel que, pour tout $y \in \Delta(J)$,

$$\langle z' - z, G(x,y) \rangle \geqslant \langle z' - z, z' \rangle,$$

soit

$$\langle z' - z, z' - G(x,y) \rangle \leqslant 0.$$

(iv) signifie que tout demi-espace contenant C est approchable par le joueur 1. (iii) $\Rightarrow$ (iv) est donc clair. (iv) $\Rightarrow$ (i) est similaire à (ii) $\Rightarrow$ (i). □

Les théorèmes précédents 2.3 et 2.4 sont, à des différences mineures de formulation près, dues à Blackwell (1956). Indiquons que X. Spinat (2002) a montré récemment le résultat suivant :

Théorème 2.5. — *Un ensemble est approchable par le joueur 1 si et seulement si il contient un B-ensemble pour le joueur 1.*

Cela implique notamment qu'ajouter la condition

$$d_t \underset{t\to\infty}{\longrightarrow} 0 \quad \mathbb{P}_{\sigma,\tau}\, p.s.$$

dans la définition d'approchabilité ne change pas la notion.

2.2. Approchabilité par le joueur 1 versus évitabilité par le joueur 2. — Le résultat suivant est un corollaire du théorème 2.4, voir (ii).

Corollaire 2.6. — *Tout sous-ensemble convexe de $\mathbb{R}^K$ est soit approchable par le joueur 1, soit évitable par le joueur 2.*

Remarque 2.7. — Cas de la dimension 1. On montre que si K est un singleton, alors tout ensemble est soit approchable par le joueur 1, soit évitable par le joueur 2.

Exemple 2.8 **(Existence d'un ensemble ni approchable par le joueur 1, ni évitable par le joueur 2 en dimension 2 (voir Sorin, 2002))**

Posons

$$G = \begin{pmatrix} (0,0) & (0,0) \\ (1,0) & (1,1) \end{pmatrix},$$

$$C = \{(1/2, v) \mid 0 \leqslant v \leqslant 1/4\} \cup \{(1,v), 1/4 \leqslant v \leqslant 1\}.$$

Considérons la stratégie suivante σ du joueur 1 dans un jeu à $2T$ étapes : σ joue B(as) pendant les T premières étapes, puis de deux choses l'une : si la deuxième coordonnée des paiements est en moyenne pendant les T premières étapes supérieure (resp. strictement inférieure) à $1/2$, σ joue B (resp. Haut) les T étapes suivantes. Quelque que soit la stratégie du joueur 2, on aura $\overline{g}_{2T} \in C$. Donc C n'est pas évitable par le joueur 2.

On peut montrer que C n'est pas non plus approchable par le joueur 1 en considérant des stratégies du joueur 2 qui joue D(roite) pendant longtemps, induisant un paiement sur la première diagonale, puis G(auche) toujours.

2.3. Approchabilité faible. — On peut affaiblir la notion d'approchabilité en abandonnant l'uniformité de la stratégie par rapport au temps.

Définition 2.9

– C est faiblement approchable par le joueur 1 si :

$$\forall \varepsilon > 0,\ \exists T,\ \forall t \geqslant T,\ \exists \sigma \in \Sigma,\ \forall \tau \in \mathcal{T}, \quad \mathbb{E}_{\sigma,\tau}(d_t) \leqslant \varepsilon.$$

– C est faiblement approchable par le joueur 2 si :

$$\forall \varepsilon > 0,\ \exists T,\ \forall t \geqslant T,\ \exists \tau \in \mathcal{T},\ \forall \sigma \in \Sigma, \quad \mathbb{E}_{\sigma,\tau}(d_t) \leqslant \varepsilon.$$

– C est faiblement évitable par le joueur 1, resp. joueur 2, s'il existe $\delta > 0$ tel que $\{z \in \mathbb{R}^K \mid d(z, C) \geqslant \delta\}$ soit faiblement approchable par le joueur 1, resp. joueur 2.

L'ensemble C de l'exemple 2.8 est faiblement approchable par le joueur 1. N. Vieille (1992) a prouvé le résultat suivant *via* l'introduction de certains jeux différentiels.

Théorème 2.10. — *Tout sous-ensemble de $\mathbb{R}^K$ est faiblement approchable par le joueur 1 ou faiblement repoussable par le joueur 2.*

2.4. Retour au modèle standard à manque d'information d'un seul côté : stratégie optimale explicite pour le joueur non informé. — Revenons ici au formalisme de la section 1. On a une famille de matrices $(G^k)_{k \in K}$, et une probabilité initiale p sur $\Delta(K)$. D'après le théorème 1.15, le jeu $\Gamma_\infty(p)$ a une valeur qui vaut $\operatorname{cav} u(p)$. En considérant un hyperplan tangent à $\operatorname{cav} u$ au point p, on peut trouver un vecteur $\ell \in \mathbb{R}^K$ tel que

$$\langle \ell, p \rangle = \operatorname{cav} u(p) \quad \text{et} \quad \forall q \in \Delta(K),\ \langle \ell, q \rangle \geqslant \operatorname{cav} u(q) \geqslant u(q).$$

Considérons maintenant l'orthant $C = \{z \in \mathbb{R}^K \mid \forall\, k \in K,\, z^k \leqslant \ell^k\}$. Soit $q = (q^k)_k$ dans $\mathbb{R}^K$.

S'il existe k tel que $q^k > 0$, on a :

$$\inf_{c \in C} \langle q, c\rangle = -\infty \leqslant \max_{y \in \Delta(J)} \min_{x \in \Delta(I)} \sum_{k \in K} q^k G^k(x, y).$$

Supposons maintenant que $q^k \leqslant 0$ pour tout k, avec $q \neq 0$. Posons $q' = -q$ et $s = \sum_k q'^k$.

$$\begin{aligned}
\inf_{c \in C} \langle q, c\rangle &= \sum_{k \in K} q^k \ell^k = -s \sum_{k \in K} \frac{q'^k}{s} \ell^k = -s \langle \ell, q'/s\rangle \\
&\leqslant -s\, u(q'/s) \\
&\leqslant -s \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} \sum_{k \in K} \frac{q'^k}{s} G^k(x, y) \\
&= \max_{y \in \Delta(J)} \min_{x \in \Delta(I)} \sum_{k \in K} q^k G^k(x, y)
\end{aligned}$$

Donc l'analogue, pour le joueur 2, de la condition (iv) du théorème 2.4 est vérifiée. C est un B-ensemble pour le joueur 2, et est donc approchable par le joueur 2. Une stratégie τ adaptée à C pour le joueur 2 vérifie, par le théorème 2.3 : $\forall\, \sigma \in \Sigma$, $\forall\, k \in K$,

$$\mathbb{E}_{\sigma,\tau}\left(\frac{1}{T} \sum_{t=1}^{T} G^k(\widetilde{i}_t, \widetilde{j}_t) - \ell^k\right) \leqslant \mathbb{E}_{\sigma,\tau}\left(d\left(\frac{1}{T} \sum_{t=1}^{T} G^k(\widetilde{i}_t, \widetilde{j}_t), C\right)\right) \leqslant \frac{2M}{\sqrt{T}},$$

où M est ici un majorant de la norme euclidienne des vecteurs $(G^k(i,j))_{k \in K}$, avec $i \in I$ et $j \in J$. Donc

$$\begin{aligned}
\gamma_T^p(\sigma, \tau) &= \sum_{k \in K} p^k \left(\frac{1}{T} \sum_{t=1}^{T} \mathbb{E}_{\sigma,\tau}(G^k(\widetilde{i}_t, \widetilde{j}_t))\right) \\
&\leqslant \langle p, \ell\rangle + \frac{2M}{\sqrt{T}} = \operatorname{cav} u(p) + \frac{2M}{\sqrt{T}}.
\end{aligned}$$

La stratégie « d'approchabilité » τ est donc optimale pour le joueur 2 dans $\Gamma_\infty(p)$.

3. Manque d'information des deux côtés

On ne suppose plus ici qu'un des deux joueurs a toute l'information. On se donne un ensemble produit $K \times L$ et une famille $(G^{k,l})_{(k,l) \in K \times L}$

de matrices de même taille $I \times J$, ainsi que des probabilités initiales p sur K et q sur L. K, L, I et J sont supposés finis non vides. Le jeu $\Gamma_\infty(p,q)$ est le suivant :

– un état de la nature (k,l) est tiré selon la probabilité produit $p \otimes q$, puis k est annoncé uniquement au joueur 1 et l est annoncé uniquement au joueur 2.

– les joueurs répètent ensuite le jeu matriciel $G^{k,l}$, et observent après chaque étape les actions jouées dans $I \times J$.

k (resp. l) représente donc l'information initiale du joueur 1 (resp. du joueur 2). On s'est placé pour simplifier dans le cas d'informations initiales indépendantes, voir la remarque 3.6 en fin de section pour plus de généralité. Le modèle de manque d'information des deux côtés fut également introduit par Aumann et Maschler dans les années soixante (réédition en 1995). On peut consulter avec profit Zamir, 1992, et Sorin, 2002.

Une stratégie du joueur 1 est un élément $\sigma = (\sigma_t)_{t \geqslant 1}$, avec $\sigma_t : K \times (I \times J)^{t-1} \to \Delta(I)$ pour tout t. Une stratégie du joueur 2 est un élément $\tau = (\tau_t)_{t \geqslant 1}$, avec $\tau_t : L \times (I \times J)^{t-1} \to \Delta(J)$ pour tout t. On note respectivement Σ et $\mathcal{T}$ les ensembles de stratégies des joueurs 1 et 2. (p,q,σ,τ) induit une probabilité sur l'ensemble des parties $K \times L \times (I \times J)^\infty$ muni de la tribu produit, et pour tout entier strictement positif T on pose :

$$\gamma_T^{p,q}(\sigma,\tau) = \mathbb{E}_{\sigma,\tau}^{p,q}\left(\frac{1}{T}\sum_{t=1}^{T} G^{\widetilde{k},\widetilde{l}}(\widetilde{\imath}_t, \widetilde{\jmath}_t)\right).$$

Là encore les « tilde » indiquent des variables aléatoires. $\gamma_T^{p,q}$ est la fonction de paiement du jeu finiment répété à T étapes. Il a une valeur notée $v_T(p,q)$. On note aussi :

$$u(p,q) = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} \sum_{k,l} p^k q^l G^{k,l}(x,y). \tag{3}$$

$u(p,q)$ est la valeur du jeu matriciel $\sum_{k,l} p^k q^l G^{k,l}$, appelé « jeu non révélateur à (p,q) », au sens où aucun des joueurs n'utilise son information.

Pour une application continue $f : \Delta(K) \times \Delta(L) \to \mathbb{R}$, on note $\mathrm{cav_I}\, f$ la concavification de f par rapport à la première variable : pour tout (p,q) dans $\Delta(K) \times \Delta(L)$, $\mathrm{cav_I}\, f(p,q)$ est la valeur en p de la plus petite fonction concave de $\Delta(K)$ dans $\mathbb{R}$ qui soit supérieure à $f(.,q)$. De même, on note $\mathrm{vex}_{\mathrm{II}}\, f$ la convexification de f par rapport à la seconde variable. On montre que $\mathrm{cav}_{\mathrm{I}}\, f$ et $\mathrm{vex}_{\mathrm{II}}\, f$ sont continues, et on peut former les composées $\mathrm{cav}_{\mathrm{I}}\, \mathrm{vex}_{\mathrm{II}}\, f$ et $\mathrm{vex}_{\mathrm{II}}\, \mathrm{cav}_{\mathrm{I}}\ f$. Ce sont des fonctions concaves par rapport à la première variable, convexes par rapport à la seconde, et on a toujours $\mathrm{cav}_{\mathrm{I}}\, \mathrm{vex}_{\mathrm{II}}\, f(p,q) \leqslant \mathrm{vex}_{\mathrm{II}}\, \mathrm{cav}_{\mathrm{I}}\ f(p,q)$.

Étant donnés (σ,τ) dans $\Sigma \times \mathcal{T}$, on définit pour toute histoire $h_t = (i_1, j_1, \ldots, i_t, j_t) \in (I \times J)^t$ les *a posteriori* :

$$p_t(\sigma,\tau,h_t) = (p_t^k(\sigma,\tau,h_t))_{k\in K} = \left(\mathbb{P}^{p,q}_{\sigma,\tau}(\widetilde{k} = k|h_t)\right)_{k\in K}.$$

$$q_t(\sigma,\tau,h_t) = (q_t^l(\sigma,\tau,h_t))_{l\in L} = \left(\mathbb{P}^{p,q}_{\sigma,\tau}(\widetilde{l} = l|h_t)\right)_{l\in L}.$$

On procède comme dans le modèle standard à manque d'information d'un seul côté (voir la partie 1.5) : $p_t(\sigma,\tau,h_t) = p_t(\sigma,h_t)$ et $q_t(\sigma,\tau,h_t) = q_t(\tau,h_t)$. Par indépendance, on a pour tout (k,l) :

$$\mathbb{P}^{p,q}_{\sigma,\tau}\left((\widetilde{k},\widetilde{l}) = (k,l)|h_t\right) = p_t^k(\sigma,h_t)q_t^l(\tau,h_t).$$

Par rapport à $\mathbb{P}^{p,q}_{\sigma,\tau}$, $(p_t(\sigma))_{t\geqslant 0}$ et $(q_t(\tau))_{t\geqslant 0}$ sont des martingales. Rappelons que l'on a pour tout T, comme dans la preuve du lemme 1.12.

$$\mathbb{E}^{p,q}_{\sigma,\tau}\left(\sum_{t=0}^{T-1}(p_{t+1}^k(\sigma) - p_t^k(\sigma))^2\right) = \mathbb{E}^{p,q}_{\sigma,\tau}\left((p_T^k(\sigma))^2\right) - (p^k)^2 \leqslant p^k(1-p^k).$$

Définition 3.1. — Le contenu informationnel d'une stratégie σ du joueur 1 est :

$$I(\sigma) = \sup_{\tau\in\mathcal{T}} \mathbb{E}^{p,q}_{\sigma,\tau}\left(\sum_{k\in K}\sum_{t=0}^{\infty}\left(p_{t+1}^k(\sigma) - p_t^k(\sigma)\right)^2\right).$$

Par linéarité de l'espérance par rapport à τ, le sup peut être pris sur les stratégies τ du joueur 2 qui sont à la fois pures et indépendantes de l'état $l \in L$.

3.1. Maxmin et Minmax de $\Gamma_\infty(p,q)$.

— Rappelons par définition que le joueur 1 garantit le réel v si :

$$\forall\, \varepsilon > 0,\ \exists\, \sigma,\ \exists\, T,\ \forall\, t \geqslant T,\ \forall\, \tau,\quad \gamma_t^{p,q}(\sigma,\tau) \geqslant v - \varepsilon.$$

On dit que le joueur 2 défend v si :

$$\forall\, \varepsilon > 0,\ \forall\, \sigma,\ \exists\, T,\ \exists\, \tau,\ \forall\, t \geqslant T,\quad \gamma_t^{p,q}(\sigma,\tau) \leqslant v + \varepsilon.$$

Si à la fois le joueur 1 garantit v et le joueur 2 défend v, on dit que v est le maxmin du jeu répété $\Gamma_\infty(p,q)$. En particulier c'est le plus grand réel qui peut être garanti par le joueur 1. On définit de même le minmax du jeu répété : s'il existe, c'est l'unique réel qui peut être à la fois garanti par le joueur 2 et défendu par le joueur 1. Le jeu $\Gamma_\infty(p,q)$ a une valeur si et seulement si il existe un réel qui peut être garanti par les deux joueurs. C'est équivalent à l'existence et à l'égalité du maxmin et du minmax. On a ici :

***Théorème 3.2* (Aumann Maschler Stearns).** — *Dans le jeu répété $\Gamma_\infty(p,q)$, le maxmin existe et vaut* $\mathrm{cav}_{\mathrm{I}}\,\mathrm{vex}_{\mathrm{II}}\,u(p,q)$*, le minmax existe et vaut* $\mathrm{vex}_{\mathrm{II}}\,\mathrm{cav}_{\mathrm{I}}\,u(p,q)$.

Démonstration. — On se place dans le jeu répété $\Gamma_\infty(p,q)$.

(1) Tout d'abord, le joueur 2 peut jouer sans tenir compte de son information. On se trouve alors dans un jeu à manque d'information d'un seul côté comme dans la section 1. Le joueur 2 garantit donc $\mathrm{cav}\, v_q(p)$, où pour tout p' dans $\Delta(K)$,

$$v_q(p') = \max_{x\in\Delta(I)} \min_{y\in\Delta(J)} \sum_{k\in K} p'^k \Big(\sum_{l\in L} q^l G^{k,l}\Big)(x,y) = u(p',q).$$

Donc le joueur 2 garantit $\mathrm{cav}_{\mathrm{I}}\,u(p,q)$.

Le joueur 2 peut utiliser son information, et en utilisant le lemme 1.6 avec le joueur 2 comme joueur informé (qui minimise les paiements, donc en changeant cav en vex car $-\mathrm{cav}(-f) = \mathrm{vex}(f)$), on obtient que le joueur 2 garantit ici $\mathrm{vex}_{\mathrm{II}}\,\mathrm{cav}_{\mathrm{I}}\,u(p,q)$. Symétriquement, le joueur 1 garantit $\mathrm{cav}_{\mathrm{I}}\,\mathrm{vex}_{\mathrm{II}}\,u(p,q)$.

(2) Pour conclure, il suffit de montrer que le joueur 2 défend $\mathrm{cav}_{\mathrm{I}}\,\mathrm{vex}_{\mathrm{II}}\,u(p,q)$. Fixons une stratégie σ du joueur 1, et soit $\varepsilon \in (0,1)$.

Soit τ^* une stratégie du joueur 2 qui soit indépendante de l, et soit une date N tel que

$$\mathbb{E}^{p,q}_{\sigma,\tau^*}\Big(\sum_{k\in K}\sum_{t=0}^{N-1} \big(p^k_{t+1}(\sigma) - p^k_t(\sigma)\big)^2\Big) \geqslant I(\sigma) - \varepsilon.$$

On définit la stratégie τ du joueur 2 de la façon suivante. Pendant les N premières étapes, τ joue selon τ^*, de façon à extraire quasiment toute l'information contenue dans σ. Supposons que $h_N \in (I\times J)^N$ soit jouée lors des N premiers coups. Le joueur 2 calcule alors l'a posteriori $p_N(\sigma, h_N) \in \Delta(K)$, et joue une stratégie optimale dans le *jeu répété à manque d'information d'un seul côté* où :

– Le joueur 2 est informé de $l \in L$, initialement tiré selon q.

– Le joueur 1 est non informé, et si $l \in L$ est l'état sélectionné, le jeu matriciel joué à chaque étape est $\sum_{k\in K} p_N^k(\sigma, h_N) G^{k,l}$.

Par le théorème 1.15, ce jeu a une valeur qui vaut $\mathrm{vex}_{\mathrm{II}}\, u(p_N(\sigma, h_N), q)$, et les joueurs y ont des stratégies optimales. À partir de la date $N+1$, la stratégie τ joue une stratégie optimale dans ce jeu à manque d'information d'un seul côté.

Toutes les probabilités et espérances qui suivent sont prises par rapport à $\mathbb{P}^{p,q}_{\sigma,\tau}$. On a pour toute date $T \geqslant N$:

$$\begin{aligned}
\mathbb{E}\Bigg(\sum_k \big(p_T^k(\sigma) - p_N^k(\sigma)\big)^2 \Bigg) &= \mathbb{E}\Bigg(\sum_k \big(p_T^k(\sigma)\big)^2 - \big(p_N^k(\sigma)\big)^2 \Bigg) \\
&= \mathbb{E}\Bigg(\sum_k \sum_{t=N}^{T-1} \big(p_{t+1}^k(\sigma)\big)^2 - \big(p_t^k(\sigma)\big)^2 \Bigg) \\
&= \mathbb{E}\Bigg(\sum_k \sum_{t=N}^{T-1} \big(p_{t+1}^k(\sigma) - p_t^k(\sigma)\big)^2 \Bigg) \\
&\leqslant \varepsilon \qquad\qquad \text{par définition de } \tau^*.
\end{aligned}$$

L'inégalité de Cauchy-Schwartz,

$$\mathbb{E}\left(\|p_T(\sigma) - p_N(\sigma)\|_1\right) \leqslant \sqrt{|K|}\, \sqrt{\sum_k \big(\mathbb{E}(p_T^k(\sigma) - p_N^k(\sigma))\big)^2}$$

donne alors :

$$\mathbb{E}\left(\|p_T(\sigma) - p_N(\sigma)\|_1\right) \leqslant \sqrt{|K|}\sqrt{\varepsilon} \tag{4}$$

et l'erreur commise en supposant que l'information sur $\widetilde{k}$ n'évolue plus après l'étape N est faible.

Calculons maintenant les paiements. Soit $t \geqslant N$, et $h_t \in (I \times J)^t$.

$$\begin{aligned}\mathbb{E}\big(G^{\widetilde{k},\widetilde{l}}(\widetilde{i}_{t+1},\widetilde{j}_{t+1}) \mid h_t\big) \\ &= \sum_{k,l} p_t^k(\sigma,h_t) q_t^l(\tau,h_t) G^{k,l}\left(\sigma_{t+1}(k,h_t),\tau_{t+1}(l,h_t)\right) \\ &\leqslant \sum_{k,l} p_t^k(\sigma,h_t) q_t^l(\tau,h_t) G^{k,l}\left(\overline{\sigma}(h_t),\tau_{t+1}(l,h_t)\right) \\ &\quad + \sum_k \sum_l p_t^k(\sigma,h_t) q_t^l(\tau,h_t) M \|\sigma_{t+1}(k,h_t) - \overline{\sigma}(h_t)\| \\ &= \sum_{k,l} p_t^k(\sigma,h_t) q_t^l(\tau,h_t) G^{k,l}\left(\overline{\sigma}(h_t),\tau_{t+1}(l,h_t)\right) \\ &\qquad\qquad + M\mathbb{E}\left(\|p_{t+1}(\sigma) - p_t(\sigma)\| | h_t\right)\end{aligned}$$

où comme dans le cas de manque d'information d'un seul côté, on pose $\overline{\sigma}(h_t) = \sum_{k\in K} p_t^k(\sigma,h_t)\sigma_{t+1}(k,h_t)$, la constante M majore tous les paiements en valeur absolue et on peut appliquer le lemme 1.13. On introduit maintenant $p_N = p_N(\sigma,h_N)$, où $h_N \in (I\times J)^N$ correspond aux premières étapes de h_t. Notons

$$\xi_{t+1}(\sigma,\tau)(h_t) = \sum_k \sum_l p_N^k(\sigma,h_N) q_t^l(\tau,h_t) G^{k,l}\left(\overline{\sigma}(h_t),\tau_{t+1}(l,h_t)\right).$$

Alors :

$$\begin{aligned}\mathbb{E}\left(G^{\widetilde{k},\widetilde{l}}(\widetilde{i}_{t+1},\widetilde{j}_{t+1}) | h_t\right) &\leqslant \xi_{t+1}(\sigma,\tau)(h_t) + M\|p_N(\sigma,h_N) - p_t(\sigma,h_t)\| \\ &\quad + M\mathbb{E}\left(\|p_{t+1}(\sigma) - p_t(\sigma)\| | h_t\right)\end{aligned}$$

Or τ joue indépendamment de l pendant les N premières étapes. Donc $q_N(\tau,h_N) = q$, et par le théorème 1.15 on peut trouver T_0 tel que pour tout $T \geqslant T_0$:

$$\mathbb{E}\left(\frac{1}{T}\sum_{t=N}^{N+T-1} \xi_{t+1}(\sigma,\tau)(h_t) | h_N\right) \leqslant \operatorname{vex}_{\mathrm{II}} u(p_N(\sigma,h_N),q) + \varepsilon.$$

(on pourrait même avoir

$$\mathbb{E}\left(\frac{1}{T}\sum_{t=N}^{N+T-1} \xi_{t+1}(\sigma,\tau)(h_t) | h_N\right) \leqslant \operatorname{vex}_{\mathrm{II}} u(p_N(\sigma,h_N),q)$$

pour tout T en reprenant les preuves de la sous-section 1.3). Nous obtenons ici :

$$
\begin{aligned}
&\mathbb{E}\biggl(\frac{1}{T}\sum_{t=N}^{N+T-1} G^{\widetilde{k},\widetilde{l}}(\widetilde{i}_{t+1},\widetilde{j}_{t+1})|h_N\biggr)\\
&\qquad\leqslant \operatorname{vex}_{\mathrm{II}} u(p_N(\sigma,h_N),q)+\varepsilon+\frac{M}{T}\sum_{t=N}^{N+T-1}\mathbb{E}\left(\|p_N(\sigma,h_N)-p_t(\sigma)\||h_N\right)\\
&\qquad\qquad\qquad\qquad\qquad +\frac{M}{T}\sum_{t=N}^{N+T-1}\mathbb{E}\left(\|p_{t+1}(\sigma)-p_t(\sigma)\||h_N\right),\\
&\qquad\leqslant \operatorname{vex}_{\mathrm{II}} u(p_N(\sigma,h_N),q)+\varepsilon+\frac{M}{T}\sum_{t=N}^{N+T-1}\mathbb{E}\left(\|p_N(\sigma,h_N)-p_t(\sigma)\||h_N\right)\\
&\qquad\qquad\qquad\qquad\qquad +\frac{M}{\sqrt{T}}\sum_{k}\sqrt{p_N^k(1-p_N^k)},
\end{aligned}
$$

par le lemme 1.12. Comme

$$\operatorname{vex}_{\mathrm{II}} u(p_N(\sigma,h_N),q)\leqslant \operatorname{cav}_{\mathrm{I}}\operatorname{vex}_{\mathrm{II}} u(p_N(\sigma,h_N),q),$$

l'inégalité de Jensen ainsi que l'inégalité précédente (4) donnent finalement :

$$
\begin{aligned}
&\mathbb{E}\biggl(\frac{1}{T}\sum_{t=N}^{N+T-1} G^{\widetilde{k},\widetilde{l}}(\widetilde{i}_{t+1},\widetilde{j}_{t+1})\biggr)\\
&\qquad\qquad\leqslant \operatorname{cav}_{\mathrm{I}}\operatorname{vex}_{\mathrm{II}} u(p,q)+\varepsilon+M\sqrt{|K|}\sqrt{\varepsilon}+\frac{M|K|}{\sqrt{T}},
\end{aligned}
$$

et donc pour T assez grand,

$$\mathbb{E}\biggl(\frac{1}{T}\sum_{t=1}^{T} G^{\widetilde{k},\widetilde{l}}(\widetilde{i}_t,\widetilde{j}_t)\biggr)\leqslant \operatorname{cav}_{\mathrm{I}}\operatorname{vex}_{\mathrm{II}} u(p,q)+(2+M)\sqrt{|K|}\sqrt{\varepsilon}.\quad\square$$

On a toujours $\operatorname{cav}_{\mathrm{I}}\operatorname{vex}_{\mathrm{II}} u(p,q)\leqslant \operatorname{vex}_{\mathrm{II}}\operatorname{cav}_{\mathrm{I}} u(p,q)$, et un corollaire immédiat du théorème 3.2 est que le jeu $\Gamma_\infty(p,q)$ a une valeur si et seulement si $\operatorname{cav}_{\mathrm{I}}\operatorname{vex}_{\mathrm{II}} u(p,q)\,\operatorname{vex}_{\mathrm{II}}\operatorname{cav}_{\mathrm{I}} u(p,q)$. Ce n'est pas toujours le cas, et il y a des contre-exemples à l'existence de la valeur (les premiers étant dûs à Aumann et Maschler).

Exemple 3.3. — Prenons $K = \{a, a'\}$, et $L = \{b, b'\}$, avec p et q uniformes.

$$G^{a,b} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 \end{pmatrix} \qquad G^{a,b'} = \begin{pmatrix} 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$G^{a',b} = \begin{pmatrix} -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad G^{a',b'} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

Mertens et Zamir (1971) ont montré qu'ici

$$\mathrm{cav}_{\mathrm{I}}\, \mathrm{vex}_{\mathrm{II}}\, u(p,q) = -\frac{1}{4} < 0 = \mathrm{vex}_{\mathrm{II}}\, \mathrm{cav}_{\mathrm{I}}\, u(p,q).$$

3.2. Les jeux compacts. — v_T et v_λ sont des applications M-Lipschitz, concaves par rapport à la première variable et convexes par rapport à la seconde. Rappelons que u est la valeur du jeu non révélateur. Le résultat suivant, que nous nous contentons d'énoncer, est dû à Mertens et Zamir (1971). On se place dans l'ensemble $\mathcal{C}$ des applications continues de $\Delta(K) \times \Delta(L)$ dans $\mathbb{R}$.

Théorème 3.4. — *$(v_T)_T$ et $(v_\lambda)_\lambda$ convergent uniformément vers l'unique solution f du système suivant :*

$$\begin{cases} f = \mathrm{vex}_{\mathrm{II}} \max\{u, f\} \\ f = \mathrm{cav}_{\mathrm{I}} \min\{u, f\} \end{cases}$$

L'étude du système ci-dessus peut se faire sans référence aux jeux répétés (voir Mertens et Zamir 1977, Sorin 1984, Laraki 2001a, Laraki 2001b).

Remarque 3.5. — Structure des fonctions non révélatrices u. Au sein de $\mathcal{C}$, on considère le sous-ensemble U constitué des valeurs d'une famille de matrices, au sens de l'équation (3). Il est possible de montrer que U est un sous-espace vectoriel de $\mathcal{C}$, contient les fonctions affines, est stable par produit, par passage au sup et à l'inf, et est dense dans $\mathcal{C}$ pour la topologie de la convergence uniforme.

Remarque 3.6. — Le modèle de manque d'information des deux côtés se généralise au cas où les informations initiales des joueurs ne sont plus indépendantes. Notons le nouvel ensemble d'états R (au lieu de $K \times L$ précédemment). L'état r dans R est tiré selon une probabilité connue $p = (p^r)_{r \in R}$, puis chaque joueur observe un signal déterministe dépendant de r. Cela revient à considérer, pour chaque joueur i, une partition R^i de R et à supposer que le joueur i observe de façon privée l'élément de sa partition qui contient l'état sélectionné.

À la première étape, le joueur 1 va jouer une action $x = (x^r)_{r\in R}$ mesurable par rapport à R^1, *i.e.* telle que $(r \to x^r)$ soit constante sur chaque élément de R^1. Après l'observation du premier coup du joueur 1 dans I, on peut calculer la probabilité conditionnelle sur R. Celle-ci appartiendra à l'ensemble suivant :

$$\Pi^{\mathrm{I}}(p) = \Big\{(\alpha^r p^r)_{r\in R} \mid \forall\, r\ \alpha^r \geqslant 0,\\ \textstyle\sum_r \alpha^r p^r = 1 \text{ et } (\alpha^r)_r \text{ est } R^1\text{-mesurable}\Big\}.$$

$\Pi^{\mathrm{I}}(p)$ contient p, est convexe compact dans $\Delta(R)$, et on dit qu'une application f de $\Delta(R)$ dans $\mathbb{R}$ est I-concave si pour tout p dans $\Delta(R)$ sa restriction à $\Pi^{\mathrm{I}}(p)$ est concave. Pour $g : \Delta(R) \to \mathbb{R}$ majorée, on définit $\mathrm{cav}_{\mathrm{I}}\, g$ comme la plus petite fonction I-concave supérieure à g. On définit de façon analogue l'ensemble $\Pi^{\mathrm{II}}(p)$, la notion de II- convexité et la II-convexifiée $\mathrm{vex}_{\mathrm{II}}\, g$. Ces définitions généralisent celles du cas d'informations initiales indépendantes, et les résultats des théorèmes 3.2 et 3.4 s'étendent parfaitement (voir Mertens et Zamir, 1971).

4. Somme non nulle et manque d'information d'un seul côté

Ici, deux joueurs vont répéter indéfiniment un même jeu *bimatriciel* tiré au départ selon une probabilité connue, seul le joueur 1 prenant connaissance de la réalisation du tirage. Formellement, on se donne deux familles $(A^k)_{k\in K}$ et $(B^k)_{k\in K}$ de matrices de même taille $I \times J$, et une probabilité p sur K. K, I, J, sont des ensembles finis non vides. On suppose que chaque joueur a au moins deux actions, $|I| \geqslant 2$ et $|J| \geqslant 2$, et on suppose $p^k > 0$ pour tout k de K. Le jeu $\Gamma_\infty(p)$ se déroule ainsi :

– initialement, un état de la nature k est tiré, une fois pour toutes, selon p. Le joueur 1 apprend k, pas le joueur 2.

– à chaque étape $t = 1, 2, \ldots$, simultanément le joueur 1 choisit une action $i_t \in I$ et le joueur 2 choisit une action j_t dans J. Le paiement d'étape du joueur 1 est alors $A^k(i_t, j_t)$, celui du joueur 2 est $B^k(i_t, j_t)$, et tout ce que les joueurs apprennent avant de passer à l'étape $t + 1$ est le couple (i_t, j_t).

Lorsque $B^k = -A^k$ pour tout k, le jeu est à somme nulle et on est dans le cadre de la section 1. Les ensembles de stratégies des joueurs sont définis comme dans cette section, et une paire de stratégies (σ, τ) dans $\Sigma \times \mathcal{T}$ induit une probabilité sur l'ensemble des parties $K \times (I \times J)^\infty$ muni de la tribu produit. Les paiements moyens espérés sont notés :

$$\alpha_T^p(\sigma,\tau) = \mathbb{E}_{p,\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} A^{\widetilde{k}}(\widetilde{\imath}_t, \widetilde{\jmath}_t)\right) = \sum_{k\in K} p^k \alpha_T^k(\sigma,\tau),$$

$$\beta_T^p(\sigma,\tau) = \mathbb{E}_{p,\sigma,\tau}\left(\frac{1}{T}\sum_{t=1}^{T} B^{\widetilde{k}}(\widetilde{\imath}_t, \widetilde{\jmath}_t)\right) = \sum_{k\in K} p^k \beta_T^k(\sigma,\tau).$$

Il est pratique d'utiliser ici la définition suivante.

Définition 4.1. — (σ^*, τ^*) est un équilibre du jeu répété à information incomplète $\Gamma_\infty(p)$ si :

(i) pour tout $k \in K$, $(\alpha_T^k(\sigma^*,\tau^*))_T$ et $(\beta_T^p(\sigma^*,\tau^*))_T$ convergent vers des limites respectivement notées $\alpha^k(\sigma^*,\tau^*)$ et $\beta^p(\sigma^*,\tau^*)$,

(ii)

$$\forall\,\varepsilon > 0,\ \exists\,T_0,\ \forall\,T \geqslant T_0,\ \forall\,k \in K,\ \forall\,\sigma \in \Sigma,\quad \alpha_T^k(\sigma,\tau^*) \leqslant \alpha_T^k(\sigma^*,\tau^*)+\varepsilon$$

et

$$\forall\,\varepsilon > 0,\ \exists\,T_0,\ \forall\,T \geqslant T_0,\ \forall\,\tau \in \mathcal{T},\quad \beta_T^p(\sigma^*,\tau) \leqslant \beta_T^p(\sigma^*,\tau^*) + \varepsilon.$$

$\big((\alpha^k(\sigma^*,\tau^*))_{k\in K}, (\beta^p(\sigma^*,\tau^*))\big) \in \mathbb{R}^K \times \mathbb{R}$ est alors appelé paiement d'équilibre de $\Gamma_\infty(p)$.

Comme $p \in \operatorname{int}(\Delta(K))$, la première ligne de (ii) équivaut à :

$$\forall\,\varepsilon > 0,\ \exists\,T_0,\ \forall\,T \geqslant T_0,\ \forall\,\sigma \in \Sigma,\quad \alpha_T^p(\sigma,\tau^*) \leqslant \alpha_T^p(\sigma^*,\tau^*) + \varepsilon.$$

Remarquons que cette définition d'équilibre est tout de même légèrement plus forte que la définition usuelle d'équilibre uniforme (voir le texte de T. Tomala sur les jeux répétés dans ce volume) : il est ici pratique d'imposer la convergence de $(\alpha_T^k(\sigma^*,\tau^*))_T$ pour chaque valeur de k, et non pas seulement la convergence de $(\alpha_T^p(\sigma^*,\tau^*))_T$. En quelque sorte, le joueur 1 est vu comme $|K|$ différents types possibles ayant chacun une fonction de paiement spécifique, et on veut que le paiement de chaque type converge. Si l'on est dans le cas de la somme nulle ($A^k = -B^k$ pour tout k), l'existence d'un tel équilibre implique l'existence de la valeur et de stratégies optimales pour chaque joueur.

La question de l'existence d'équilibre a été posée par Aumann, Maschler et Stearns dans les années soixante. Sorin (1983) a prouvé l'existence pour deux états de la nature, et le cas général a été résolu en 1995 par Simon, Spież et Toruńczyk (voir les théorèmes suivants 4.6 et 4.7). Hart (1985) a donné une caractérisation des paiements d'équilibres, qui n'a toutefois pas entraîné de preuve d'existence. On peut consulter Forges (1992) pour un survey.

On note, pour toute probabilité q dans $\Delta(K)$,

$$A(q) = \sum_k q^k A^k, \qquad u(q) = \max_{x \in \Delta(I)} \min_{y \in \Delta(J)} A(q)(x,y),$$

$$B(q) = \sum_k q^k B^k, \qquad v(q) = \max_{y \in \Delta(J)} \min_{x \in \Delta(I)} B(q)(x,y).$$

Si $\gamma = (\gamma(i,j))_{(i,j) \in I \times J} \in \Delta(I \times J)$,

$$A(q)(\gamma) = \sum_{(i,j) \in I \times J} \gamma(i,j) A(q)(i,j)$$

et de même on pose

$$B(q)(\gamma) = \sum_{(i,j) \in I \times J} \gamma(i,j) B(q)(i,j).$$

4.1. Existence d'un équilibre. — Exactement comme dans le cas de la somme nulle, une paire de stratégies (σ, τ) induit une suite d'*a posteriori* $(p_t(\sigma))_{t \geqslant 0}$ qui est une $\mathbb{P}_{p,\sigma,\tau}$- martingale à valeurs dans $\Delta(K)$. Du point de vue de l'existence d'un équilibre, on va se restreindre aux cas où cette martingale « bouge » au plus une fois.

Définition 4.2. — Un plan joint dans $\Gamma_\infty(p)$ est un triplet (S, λ, γ), où :

– S est un ensemble (« de messages ») fini non vide,

– $\lambda = (\lambda^k)_{k \in K}$ (« stratégie de signalling ») avec pour tout k, $\lambda^k \in \Delta(S)$ et pour tout s, $\lambda_s = \sum_{k \in K} p^k \lambda^k_s > 0$,

– $\gamma = (\gamma_s)_{s \in S}$ (« contrat ») avec pour tout s, $\gamma_s \in \Delta(I \times J)$.

L'idée, due à Aumann, Maschler et Stearns, est la suivante. Le joueur 1 observe k, puis choisit $s \in S$ selon λ^k et annonce s au joueur 2. Ensuite les joueurs jouent des actions pures correspondant aux fréquences $\gamma_s(i,j)$, pour i dans I et j dans J. Étant donné un plan joint (S, λ, γ), on définit :

– $\forall\, s \in S$, $p_s = (p_s^k)_{k\in K} \in \Delta(K)$, avec $p_s^k = p^k\lambda_s^k/\lambda_s$ pour tout k. p_s est l'*a posteriori* sur K sachant s.

– $\varphi = (\varphi^k)_{k\in K} \in \mathbb{R}^K$, avec pour tout k, $\varphi^k = \max_{s\in S} A^k(\gamma_s)$.

– $\forall\, s \in S$, $\psi_s = B(p_s)(\gamma_s)$ et

$$\psi = \sum_{k\in K} p^k \sum_{s\in S} \lambda_s^k B^k(\gamma_s) = \sum_{s\in S} \lambda_s \psi_s.$$

Définition 4.3. — Un plan joint (S, λ, γ) est un *plan joint si :

(i) $\forall\, s \in S$, $\psi_s \geqslant \operatorname{vex} v(p_s)$.

(ii) $\forall\, k \in K$, $\forall\, s \in S$ tel que $p_s^k > 0$, $A^k(\gamma_s) = \varphi^k$ (« incitation du joueur 1 à choisir s selon λ^k »).

(iii) $\forall\, q \in \Delta(K)$, $\langle \varphi, q\rangle \geqslant u(q)$.

Étant donné un *plan joint, on définit un couple de stratégies (σ^*, τ^*) adapté au plan joint. Pour tout message s de S, fixons tout d'abord une suite $(i_t^s, j_t^s)_{t\geqslant 1}$ d'éléments de $I \times J$ telle que pour tout couple (i, j), la suite des fréquences empiriques converge vers la probabilité correspondante :

$$\frac{1}{T}\big|\{t \mid 1 \leqslant t \leqslant T,\ (i_t^s, j_t^s) = (i_t, j_t)\}\big| \underset{T\to\infty}{\longrightarrow} \gamma_s(i, j).$$

On se donne également un entier ℓ et une application injective $f : S \to I^\ell$ correspondant à un code entre les joueurs pour annoncer un élément de S.

On définit σ^* de la façon suivante. Le joueur 1 observe l'état k sélectionné, puis choisit s selon la probabilité λ^k, et annonce s au joueur 2 en jouant $f(s)$ pendant les ℓ premiers coups. Enfin, σ^* joue i_t^s à chaque date $t > \ell$ tant que le joueur 2 joue j_t^s. Si à une date $t > \ell$ le joueur 2 ne joue pas j_t^s alors le joueur 1 se met à jouer une stratégie de punition du joueur 2 dans le jeu de probabilité initiale p_s, *i.e.* le joueur 1 joue une stratégie optimale dans le jeu à somme nulle de probabilité initiale p_s où les paiements du joueur 1 sont $(-B^k)_{k\in K}$.

On définit maintenant τ^*. Le joueur 2 joue arbitrairement au début puis à la fin de l'étape ℓ il déduit le message s des coups du joueur 1. Il joue ensuite à chaque date $t > \ell$ l'action j_t^s tant que le joueur 1 joue i_t^s. Si à une date $t > \ell$, le joueur 1 ne joue pas i_t^s, ou si les ℓ premiers coups du joueur 1 ne correspondent à aucun message, alors le joueur 2 se met à jouer une stratégie de punition $\overline{\tau}$ telle que :

$$\forall\, \varepsilon > 0,\ \exists\, T_0,\ \forall\, T \geqslant T_0,\ \forall\, \sigma \in \Sigma,\ \forall\, k \in K, \quad \alpha_T^k(\sigma, \overline{\tau}) \leqslant \varphi^k + \varepsilon.$$

Une telle stratégie $\overline{\tau}$ existe en raison de (iii) : c'est une stratégie d'approchabilité par le joueur 2 de l'orthant

$$\{x \in \mathbb{R}^K \mid \forall\, k \in K,\ x^k \leqslant \varphi^k\}$$

(voir section 2, 2.4).

***Lemme 4.4* (Sorin, 1983).** — *Un couple de stratégies (σ^*, τ^*) adapté à un *plan joint est un équilibre de $\Gamma_\infty(p)$.*

Démonstration. — Pour tout k,

$$\alpha^k(\sigma^*, \tau^*) = \sum_{s\in S} \lambda_s^k A^k(\gamma_s) = \sum_{s\in S} \lambda_s^k \varphi^k \varphi^k,$$

d'après (ii), et

$$\beta(\sigma^*, \tau^*) = \sum_{k\in K} p^k \sum_{s\in S} \lambda_s^k B^k(\gamma_s) = \psi.$$

Supposons que le joueur 2 joue τ^*. L'existence de $\overline{\tau}$ fait qu'aucune déviation détectable du joueur 1 n'est profitable, et donc que si l'état est k, le joueur 1 ne pourra gagner plus que $\max_{s'\in S} A^k(\gamma_{s'})$. Or ceci vaut $\varphi^k = \alpha^k(\sigma^*, \tau^*)$. La preuve peut être uniforme en σ et on obtient :

$$\forall\, \varepsilon > 0,\ \exists\, T_0,\ \forall\, T \geqslant T_0,\ \forall\, k \in K,\ \forall\, \sigma \in \Sigma, \quad \alpha_T^k(\sigma, \tau^*) \leqslant \alpha^k(\sigma^*, \tau^*) + \varepsilon.$$

Supposons enfin que le joueur 1 joue σ^*. La condition (i) implique que le joueur 2 qui joue τ^* gagne, si le message vaut s, au moins $\operatorname{vex} v(p_s)$. Comme $\operatorname{vex} v(p_s)$ $(= -\operatorname{cav}(-v(p_s)))$ est la valeur, pour le joueur 2 de paiements $(B^k)_k$, du jeu à somme nulle de probabilité initiale p_s, le joueur 2 craint la punition du joueur 1, et

$$\forall\, \varepsilon > 0,\ \exists\, T_0,\ \forall\, T \geqslant T_0,\ \forall\, \tau \in \mathcal{T}, \quad \beta_T^p(\sigma^*, \tau) \leqslant \sum_{s\in S} \lambda_s \psi_s + \varepsilon = \psi + \varepsilon.$$

$\square$

Afin de prouver l'existence d'équilibres dans $\Gamma_\infty(p)$, on se retrouve à chercher des *plans joints. L'idée est tout d'abord de considérer, pour chaque probabilité r sur K, l'ensemble des vecteurs φ possibles s'il y a un plan joint dont r fait partie des a posteriori. Ceci amène à considérer la correspondance[(1)] suivante.

$$\begin{aligned}\Phi : \Delta(K) &\rightrightarrows \mathbb{R}^K\\ r &\longmapsto \{(A^k(\gamma))_{k\in K} \mid \gamma \in \Delta(I\times J), B(r)(\gamma) \geqslant \operatorname{vex} v(r)\}\end{aligned}$$

Il est facile de voir que Φ a un graphe compact et des valeurs convexes non vides, et vérifie :

$$\forall\, r \in \Delta(K), \forall\, q \in \Delta(K), \exists\, \varphi \in \Phi(r), \quad \langle \varphi, q\rangle \geqslant u(q).$$

Supposons maintenant que l'on trouve un ensemble fini $(p_s)_{s\in S}$ d'éléments de $\Delta(K)$, ainsi que des vecteurs de $\mathbb{R}^K$ φ et φ_s pour tout s tels que :

- $p \in \operatorname{conv}\{p_s \mid s \in S\}$,
- $\forall\, q \in \Delta(K)$, $\langle \varphi, q\rangle \geqslant u(q)$,
- $\forall\, s \in S$, $\varphi_s \in \Phi(p_s)$,
- $\forall\, s \in S$, $\forall\, k \in K$, $\varphi_s^k \leqslant \varphi^k$ avec égalité si $p_s^k > 0$.

Alors il est facile de construire un *plan joint. On se retrouve donc à essayer de démontrer le résultat suivant :

Proposition 4.5. — *Soient $p \in \operatorname{int}(\Delta(K))$, $u : \Delta(K) \to \mathbb{R}$ une application continue, et $\Phi : \Delta(K) \rightrightarrows \mathbb{R}^K$ une correspondance de graphe compact à valeurs convexes non vides tels que :*

$$\forall\, r \in \Delta(K), \forall\, q \in \Delta(K), \exists\, \varphi \in \Phi(r), \quad \langle \varphi, q\rangle \geqslant u(q).$$

Alors il existe un ensemble fini S, une famille $(p_s)_{s\in S}$ d'éléments de $\Delta(K)$, ainsi que des vecteurs de $\mathbb{R}^K$ φ et φ_s pour tout s de S tels que :

- *$p \in \operatorname{conv}\{p_s \mid s \in S\}$,*
- *$\forall\, q \in \Delta(K)$, $\langle \varphi, q\rangle \geqslant u(q)$,*
- *$\forall\, s \in S$, $\varphi_s \in \Phi(p_s)$,*
- *$\forall\, s \in S$, $\forall\, k \in K$, $\varphi_s^k \leqslant \varphi^k$ avec égalité si $p_s^k > 0$.*

[(1)]Rappelons qu'une correspondance F d'un ensemble X dans un ensemble Y est une application de X dans l'ensemble des parties de Y. Le graphe de la correspondance F est alors définie comme $\{(x,y) \in X \times Y \mid y \in F(x)\}$.

La preuve de ce résultat repose sur un théorème de point fixe de type Borsuk-Ulam[2] démontré par Simon, Spież et Toruńczyk (1995) *via* des outils de topologie algébrique (on peut voir Renault, 2000, ou Simon 2002 pour le passage à la proposition 4.5). Une version simplifiée de ce résultat de type point fixe est donnée maintenant :

Théorème 4.6 **(Simon, Spież et Toruńczyk, 1995).** — *Soient C un compact d'un espace euclidien de dimension n, $x \in \mathrm{int}(C)$, et Y une union finie de sous-espaces affines de dimension $n-1$ d'un espace euclidien. Soit F une correspondance de C dans Y de graphe compact à valeurs convexes non vides. Alors il existe $L \subset \partial C$ et $y \in Y$ tels que :*

$$\forall\, l \in L, \quad y \in F(l) \quad \textit{et} \quad x \in \mathrm{conv}(L).$$

Remarquons que pour $n = 1$ (qui correspond à deux états possibles), l'image par F de la composante connexe de C contenant x est nécessairement un singleton, donc le résultat est clair. Tous comptes faits, on aboutit donc à :

Théorème 4.7 **(Simon, Spież et Toruńczyk, 1995).** — *Il existe un *plan joint. Donc il existe un équilibre dans le jeu répété $\Gamma_\infty(p)$.*

4.2. Caractérisation des paiements d'équilibre. — On présente ici la caractérisation des paiements d'équilibre due à S. Hart (1985). Notons dans cette partie $p_0 \in \mathrm{int}(\Delta(K))$ la probabilité initiale. Soit un équilibre (σ^*, τ^*) de $\Gamma_\infty(p_0)$ de paiement $(a, \beta) \in \mathbb{R}^K \times \mathbb{R}$. On a d'après la définition 4.1 :

$$\forall\, k \in K, \forall\, \varepsilon > 0, \exists\, T_0, \forall\, T \geqslant T_0, \forall\, \sigma \in \Sigma, \quad \alpha_T^k(\sigma, \tau^*) \leqslant a^k + \varepsilon.$$

Donc l'orthant $\{x \in \mathbb{R}^K \mid \forall\, k \in K, x^k \leqslant a^k\}$ est approchable par le joueur 2, et on montre avec le théorème 2.4 (voir aussi la partie 2.4) :

$$\forall\, q \in \Delta(K), \quad \langle a, q\rangle \geqslant u(q) \tag{5}$$

La propriété (5) s'appelle la condition de rationalité individuelle du joueur 1, et ne dépend pas de la probabilité initiale dans $\mathrm{int}(\Delta(K))$. En ce qui concerne le joueur 2, on a :

$$\forall\, \varepsilon > 0, \exists\, T_0, \forall\, T \geqslant T_0, \forall\, \tau \in \mathcal{T}, \quad \beta_T^{p_0}(\sigma^*, \tau) \leqslant \beta + \varepsilon,$$

[2] Pour toute application continue de la sphère unité de $\mathbb{R}^{n+1}$ dans $\mathbb{R}^n$ il existe deux points diamétralement opposés ayant la même image.

donc par le théorème 1.15 :

$$\beta \geqslant \operatorname{vex} v(p_0). \tag{6}$$

La propriété (6) s'appelle la condition de rationalité individuelle du joueur 2 : à l'équilibre, ce joueur doit avoir au moins la valeur du jeu où les paiements du joueur 1 sont opposés aux siens.

Supposons un instant que σ^* soit une stratégie non révélatrice du joueur 1, au sens où le joueur 1 joue indépendamment de l'état k sélectionné. Supposons également que les joueurs jouent des actions dont les fréquences empiriques convergent vers les probabilités d'une distribution $\pi = (\pi_{i,j})_{(i,j)\in I\times J} \in \Delta(I \times J)$. On aura alors :

$$\forall\, k \in K, \quad a^k \sum_{i,j} \pi_{i,j} A^k(i,j) \quad \text{et} \quad \beta = \sum_k p_0^k \sum_{i,j} \pi_{i,j} B^k(i,j),$$

et si les conditions de rationalité individuelle sont vérifiées, alors aucune déviation détectable d'un joueur n'est profitable. Ceci amène à définir l'ensemble suivant, où M est une constante fixée égale à $\max\{|A^k(i,j)|, |B^k(i,j)|, (i,j) \in I \times J\}$, et où $\mathbb{R}_M = [-M, M]$.

Définition 4.8. — Soit G l'ensemble des triplets $(a, \beta, p) \in \mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$ tels que :

(1) $\forall\, q \in \Delta(K)$, $\langle a, q\rangle \geqslant u(q)$,

(2) $\beta \geqslant \operatorname{vex} v(p)$,

(3) il existe $\pi \in \Delta(I \times J)$ tel que $\beta = \sum_k p^k \sum_{i,j} \pi_{i,j} B^k(i,j)$ et pour tout $k \in K$, $a^k \geqslant \sum_{i,j} \pi_{i,j} A^k(i,j)$ avec égalité si $p^k > 0$.

On est amené à considérer toutes les probabilités initiales possibles $p \in \Delta(K)$ car la variable d'état importante du modèle est, là encore, la martingale des *a posteriori* du joueur 2 sur l'état de la nature. Pour $p \in \operatorname{int}(\Delta(K))$, $\{(a, \beta) \mid (a, \beta, p) \in G\}$ est l'ensemble des paiements d'équilibre non révélateurs de $\Gamma_\infty(p)$. La définition suivante est essentielle.

Définition 4.9. — On définit l'ensemble G^* comme l'ensemble des éléments $g = (a, \beta, p) \in \mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$ tels qu'il existe un espace probabilisé $(\Omega, \mathcal{A}, Q)$, une suite croissante $(\mathcal{F}_n)_{n\geqslant 1}$ de sous-tribus finies de $\mathcal{A}$, une suite de v.a. $(g_n)_{n\geqslant 1}(a_n, \beta_n, p_n)_{n\geqslant 1}$ définies sur $(\Omega, \mathcal{A})$ à valeurs dans $\mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$ satisfaisant :

(i) $g_1 = (a, \beta, p)$ p.s.,

(ii) $(g_n)_{n\geqslant 1}$ est une $(\mathcal{F}_n)_{n\geqslant 1}$ martingale,
(iii) $\forall\, n \geqslant 1$, $a_{n+1} = a_n$ p.s. ou $p_{n+1} = p_n$ p.s., et
(iv) $(g_n)_n$ converge p.s. vers une v.a. g_∞ à valeurs dans G.

Oublions dans un premier temps la composante du paiement du joueur 2. Un processus $(g_n)_n$ vérifiant (ii) et (iii) s'appelle une bi-martingale, c'est une martingale telle qu'à chaque étape, il existe une des deux composantes qui n'évolue p.s. pas. G^* peut donc se voir comme l'ensemble des points de départ des bi-martingales qui convergent dans G. L'importance ici de l'ensemble G^* vient du résultat suivant.

Théorème 4.10 **(Hart, 1985).** — *Soit* $(a, \beta) \in \mathbb{R}^K \times \mathbb{R}$.

(a, β) *est un paiement d'équilibre de* $\Gamma_\infty(p_0) \Longleftrightarrow (a, \beta, p_0) \in G^*$.

Donnons maintenant dans les deux paragraphes suivants, non pas une démonstration, mais une idée approximative de la preuve du théorème 4.10.

Commençons par l'implication $\Rightarrow$. Fixons un équilibre (σ^*, τ^*) de $\Gamma_\infty(p)$. La suite des *a posteriori* $(p_t(\sigma))_{t\geqslant 0}$ est une $\mathbb{P}_{p,\sigma,\tau}$- martingale. Modifions légèrement la structure temporelle de telle sorte qu'à chaque étape le joueur 1 joue d'abord, puis que le joueur 2 joue sans avoir pris connaissance du coup du joueur 1. À chaque demi-étape où le joueur 2 joue, l'*a posteriori* reste constant. À chaque demi-étape où le joueur 1 joue, l'espérance du paiement futur du joueur 1 (qui reste à définir proprement, à l'aide notamment d'une limite de Banach) reste constante. D'où, de façon heuristique, l'apparition de la bimartingale. Enfin, par convergence des martingales bornées, au bout d'un moment tout sera fixé et on jouera alors approximativement un équilibre non révélateur pour un *a posteriori* limite, donc on convergera vers des éléments de G.

Passons maintenant à l'implication $\Leftarrow$. Soit (a, β) tel que $(a, \beta, p_0) \in G^*$, et supposons pour simplifier que la bi-martingale associée (a_n, β_n, p_n) converge en un nombre fixé N d'étapes :

$$\forall\, n \geqslant N, \quad (a_n, \beta_n, p_n) = (a_N, \beta_N, p_N) \in G.$$

On peut construire un équilibre (σ^*, τ^*) de $\Gamma_\infty(p_0)$ de paiement (a, β) de la façon suivante. À chaque fois, (a_n, β_n) sera un paiement d'équilibre du jeu de probabilité initiale p_n. Après une certaine étape, le joueur 1 jouera de façon indépendante de l'état de la nature, l'*a posteriori* du joueur 2 sera p_N, et on jouera jusqu'à la fin des temps un équilibre non révélateur de $\Gamma_\infty(p_N)$ de paiement (a_N, β_N). Comment arriver jusque là ? Souvenons-nous que, de par la définition même des paiements dans le jeu infiniment répété, le poids d'un nombre fini d'étapes est nul. Les joueurs peuvent donc passer un grand nombre d'étapes à communiquer, sans que cela n'influe sur les paiements. Soit un indice $n < N$ tel que $a_{n+1} = a_n$. Pour passer de (a_n, β_n, p_n) à $(a_n, \beta_{n+1}, p_{n+1})$, le joueur 1 peut utiliser le lemme 1.5 de splitting afin de « signaller » une partie de l'information au joueur 2. Soit maintenant un indice $n < N$ tel que $p_{n+1} = p_n$. On souhaite passer de (a_n, β_n, p_n) à $(a_{n+1}, \beta_{n+1}, p_n)$. Le joueur 1 va jouer indépendamment de l'état, et les deux joueurs vont agir de façon à convexifier leurs paiements futurs. Ceci peut se faire au moyen de procédures appelées *loteries conjointement contrôlées*, et introduites dès les années soixante (Aumann Maschler, 1995). L'idée est la suivante. Imaginons que les deux joueurs doivent décider avec probabilités égales de jouer l'équilibre $E1$ de paiement (a^1, β^1) ou de jouer l'équilibre $E2$ de paiement (a^2, β^2). Les joueurs ne sont pas nécessairement indifférents entre les deux équilibres, il est possible par exemple que le joueur 1 préfère $E1$ alors que le joueur 2 préfère $E2$. On peut procéder ainsi. Notons i et i', respectivement j et j', deux actions différentes du joueur 1, resp. joueur 2. Simultanément et indépendamment, le joueur 1 va jouer i ou i' à probabilités égales, et le joueur 2 va jouer j ou j' à probabilités égales.

$$\begin{array}{cc} & \begin{array}{cc} j & j' \end{array} \\ \begin{array}{c} i \\ i' \end{array} & \begin{pmatrix} \times & \\ & \times \end{pmatrix} \end{array}.$$

Puis les joueurs décident conjointement de jouer $E1$ si on est sur la diagonale, *i.e.* si (i, j) ou (i', j') est joué, et de jouer $E2$ sinon. Cette procédure est robuste aux déviations unilatérales : aucun des joueurs ne peut dévier de façon à empêcher que $E1$ et $E2$ soient choisis avec probabilités égales. Plus généralement, les loteries conjointement contrôlées permettent de sélectionner une alternative au sein d'un ensemble fini selon une probabilité voulue (penser aux expansions binaires), et ceci de façon robuste aux déviations de la part d'un seul joueur. S. Hart a montré qu'en combinant des étapes de « signalling » et des loteries conjointement contrôlées, il était possible de construire un équilibre de $\Gamma_\infty(p_0)$ de paiement (a, β).

4.3. Biconvexité et bimartingales. — L'analyse précédente incite à définir et étudier certaines propriétés générales dites de biconvexité. La référence ici est l'article de Aumann et Hart (1986).

Soient X et Y des convexes compacts d'espaces euclidiens, et $(\Omega, \mathcal{F}, \mathcal{P})$ un espace probabilisé sans atome.

Définition 4.11. — Un sous-ensemble B de $X \times Y$ est biconvexe si pour tous x de X et y de Y, les sections $B_{x.} = \{y' \in Y \mid (x, y') \in B\}$ et $B_{.y} = \{x' \in X \mid (x', y) \in B\}$ sont convexes. Pour B biconvexe, une application $f : B \to \mathbb{R}$ est dite biconvexe si pour tout $(x, y) \in X \times Y$, $f(., y)$ et $f(x, .)$ sont convexes.

On a, comme pour le cas classique de convexité, que si f est biconvexe, alors pour tout réel α, l'ensemble $\{(x, y) \in B \mid f(x, y) \leqslant \alpha\}$ est un ensemble biconvexe.

Définition 4.12. — Une suite $Z_n = (X_n, Y_n)_{n\geqslant 1}$ de v.a. à valeurs dans $X \times Y$ est une bimartingale si :

(1) il existe une suite croissante $(\mathcal{F}_n)_{n\geqslant 1}$ de sous-tribus finies de $\mathcal{F}$ telle que $(Z_n)_n$ soit une $(\mathcal{F}_n)_{n\geqslant 1}$-martingale.

(2) $\forall\, n \geqslant 1$, $X_n = X_{n+1}$ p.s. ou $Y_n = Y_{n+1}$ p.s.

(3) Z_1 est p.s. constante.

$(Z_n)_{n\geqslant 1}$ étant une martingale bornée, elle converge presque sûrement vers une limite Z_∞.

Définition 4.13. — Soit A un sous-ensemble mesurable de $X \times Y$. On note A^* l'ensemble des $z \in X \times Y$ pour lesquels il existe une bimartingale $(Z_n)_{n\geqslant 1}$ avec $Z_1 = z$ p.s. et convergeant vers Z_∞ avec $Z_\infty \in A$ p.s.

On peut montrer que tout espace probabilisé $(\Omega, \mathcal{F}, \mathcal{P})$ sans atome, ou encore tout produit de convexes compacts $X \times Y$ contenant A, induisent le même ensemble A^*. On peut aussi remplacer la condition (2) par :

$$\forall\, n \geqslant 1, \quad (X_n = X_{n+1} \text{ ou } Y_n = Y_{n+1}) \text{ p.s.}$$

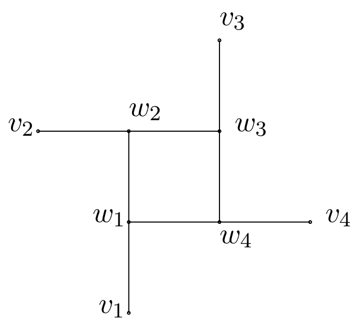
Remarquons que, sans cette condition (2) de bi-martingale, l'ensemble A^* serait seulement l'enveloppe convexe de A, et on a toujours $A \subset A^* \subset \text{conv}(A)$. Ces inclusions peuvent être strictes. Par exemple, si $X = Y = [0,1]$ et $A = \{(0,0),(1,0),(0,1)\}$, on montre que

$$A^* = \{(x,y) \in [0,1] \times [0,1] \mid x = 0 \text{ ou } y = 0\}.$$

A^* est toujours biconvexe, et contient donc biconv(A), le plus petit ensemble biconvexe qui contient A. L'inclusion biconv$(A) \subset A^*$ peut également être stricte, comme le montre l'exemple suivant :

Exemple 4.14. — On pose

$X = Y = [0,1]$,
$v_1 = (1/3,0)$, $v_2 = (0,2/3)$, $v_3 = (2/3,1)$, $v_4 = (1,1/3)$,
$w_1 = (1/3,1/3)$, $w_2 = (1/3,2/3)$, $w_3 = (2/3,2/3)$, $w_4 = (2/3,1/3)$,
$A = \{v_1,v_2,v_3,v_4\}$.



A est biconvexe, donc $A = \text{biconv}(A)$. Soit maintenant le processus markovien $(Z_n)_{n\geqslant 1}$ suivant : $Z_1 = w_1$. Si $Z_n \in A$, alors $Z_{n+1} = Z_n$. Si $Z_n = w_i$ pour un i, alors $Z_{n+1} = w_{i+1(mod\,4)}$ avec probabilité $1/2$, et $Z_{n+1} = v_i$ avec probabilité $1/2$. $(Z_n)_n$ est une bimartingale qui converge p.s. vers un point de A, donc $w_1 \in A^* \setminus \text{biconv}(A)$.

Donnons maintenant une caractérisation géométrique de l'ensemble A^*. On suppose ici A fermé. Pour chaque sous-ensemble biconvexe B de $X \times Y$ qui contient A, on note nsc(B) l'ensemble des points de B qui ne peuvent être séparés de A par une fonction biconvexe bornée continue sur A. Plus précisément,

$$\mathrm{nsc}(B) = \big\{z \in B \mid \forall f : B \to \mathbb{R} \text{ bornée biconvexe continue sur } A, \\ f(z) \leqslant \sup\{f(z'), z' \in A\}\big\}.$$

Théorème 4.15 **(Aumann et Hart, 1986).** — *A^* est le plus grand ensemble biconvexe B contenant A tel que* $\mathrm{nsc}(B) = B$.

Revenons au contexte des jeux et aux notations de la partie 4.2. Pour être précis, il faut tenir compte de la composante paiement du joueur 2, donc modifier très légèrement les définitions. G est fermé dans $\mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$. Pour $B \subset \mathbb{R}_M^K \times \mathbb{R}_M \times \Delta(K)$, B est biconvexe si pour tous a dans $\mathbb{R}_M^K$ et pour tout p dans $\Delta(K)$, les sections $\{(\beta, p'), (a, \beta, p') \in B\}$ et $\{(a', \beta), (a', \beta, p) \in B\}$ sont convexes. Une fonction réelle f définie sur un ensemble biconvexe B est dite biconvexe si pour tous a et p, $f(a, ., .)$ et $f(., ., p)$ sont convexes.

Théorème 4.16 **(Aumann et Hart, 1986).** — *G^* est le plus grand ensemble biconvexe B contenant G tel que : $\forall z \in B$, $\forall f : B \to \mathbb{R}$ biconvexe bornée continue sur A, $f(z) \leqslant \sup\{f(z'), z' \in G\}$.*

5. Extensions, divers

Concernant la modélisation, les fondements des jeux à information incomplète sont étudiés par Harsanyi (1967) et par Mertens et Zamir dans (1985). Par ailleurs, les résultats présentés précédemment ne constituent que la base des jeux répétés à information incomplète, et il existe de nombreuses extensions et variantes. La présentation suivante est sûrement imparfaite et ne prétend pas à l'exhaustivité. On n'aborde notamment pas ici les liens entre jeux répétés à information incomplète et phénomènes de réputation, « merging » de probabilités, l'apprentissage, le cheap-talk,...

5.1. Signaux. — Les modèles définis dans ce texte se généralisent au cas d'observation imparfaite. On se donne des ensembles (finis) de signaux U et V, et une application $\ell : K \times I \times J \to \Delta(U \times V)$. Après chaque étape t, si l'état est k et que (i_t, j_t) a été joué, on tire (u_t, v_t) selon $\ell(k, i_t, j_t)$. Le joueur 1, resp. joueur 2, apprend alors uniquement u_t, resp. v_t, avant de passer à l'étape $t+1$. Quand ℓ ne dépend pas de l'état k on dit que les signaux sont indépendants de l'état. Le cas où $U = V = I \times J$ et $\ell(k,i,j)$ est la mesure de Dirac sur $((i,j),(i,j))$ pour tous k, i, j, correspond aux modèles précédents dits d'observation parfaite.

Aumann et Maschler ont généralisé le théorème 1.15 au cadre général de signaux. Pour une action $x \in \Delta(I)$, une action j dans J et un état k, notons xQ_j^k la marginale sur V de la distribution $\sum_{i \in I} x_i \;\; \ell(k,i,j)$ des signaux reçus par le joueur 2. On définit l'ensemble des stratégies non révélatrices du joueur 1 comme :

$$\begin{aligned}&\mathrm{NR}(p) = \{x = (x^k)_{k\in K} \in \Delta(I)^K \mid \\ &\qquad \forall\, k \in K,\, \forall\, k' \in K \text{ tels que } p^k p^{k'} > 0,\, \forall\, j \in J,\, x^k Q_j^k = x^{k'} Q_j^{k'}\}.\end{aligned}$$

Si la probabilité initiale est p et que le joueur 1 joue selon une stratégie dans $\mathrm{NR}(p)$, l'*a posteriori* du joueur 2 restera presque sûrement égal à l'*a priori* p. La valeur du jeu NR à p devient ici :

$$\begin{aligned}u(p) &= \max_{x\in \mathrm{NR}(p)} \min_{y \in \Delta(J)} \sum_{k\in K} p^k G^k(x^k, y) \\ &= \min_{y \in \Delta(J)} \max_{x\in \mathrm{NR}(p)} \sum_{k\in K} p^k G^k(x^k, y),\end{aligned}$$

avec la convention $u(p) = -\infty$ si $\mathrm{NR}(p) = \varnothing$. Avec ces notations, on a le même énoncé que le théorème 1.15 : la valeur du jeu répété de probabilité initiale p existe et vaut $\operatorname{cav} u(p)$. Kohlberg (1975) et Mertens, Sorin et Zamir (1994) (partie B, chapitre V, 3.d.) ont généralisé dans ce cadre la construction d'une stratégie optimale explicite de type approchabilité pour le joueur 2.

En ce qui concerne les jeux à somme nulle et manque d'information des deux côtés, Mertens (1972) et Mertens et Zamir (1971 et 1977), ont généralisé l'étude de la partie 3 au cas de signaux indépendants de l'état. Dans les jeux à manque d'information d'un seul côté et somme non nulle, l'existence d'équilibre a été généralisée pour des signaux indépendants de l'état dans Renault, 2000 (voir aussi Simon, Spież et Toruńczyk 2002).

5.2. Deux joueurs somme nulle. — Indiquons tout d'abord qu'il est crucial, pour la validité du théorème 1.15, que le joueur 1 connaisse l'*a priori* p du joueur 2 sur l'état de la nature (voir Sorin et Zamir 1985, pour un jeu répété à manque d'information « d'un côté et demi » sans valeur).

Dans l'exemple 1.3, Mayberry (1967) a étudié la valeur v_λ du jeu escompté et a montré que pour $2/3 < \lambda < 1$, v_λ a une dérivée discontinue en tout point rationnel p (voir aussi Zamir, 1992 ou Sorin, 2002).

Dans le cas général, de nombreux travaux étudient les liens entre les valeurs des jeux finiment répétés, ou escomptés, et obtiennent notamment des propriétés sur la convergence de $(v_T)_T$ ou de $(v_\lambda)_\lambda$: par exemple Zamir, 1971, Zamir 1973, Mertens et Zamir 1976, de Meyer 1996a, 1996b, Laraki 2001a. On trouve de nombreuses généralisations des théorèmes principaux. Laraki (2001b) étudie des jeux dits de « splitting ». B. de Meyer a introduit la notion de « jeu dual » (voir de Meyer 1996b, de Meyer 1998, Rosenberg 1998, de Meyer et Rosenberg 1999, Laraki 2002).

Donnons juste une idée de ce jeu dual dans le cadre du modèle standard de la section 1. Soit z un paramètre dans $\mathbb{R}^K$. Dans le jeu dual $\Gamma_T^*(z)$, le joueur 1 commence par choisir secrètement l'état k. Puis à chaque étape $t \leqslant T$, les joueurs choisissent classiquement des actions i_t et j_t qui sont annoncées avant de passer à l'étape suivante. Le paiement du joueur 1 est finalement $\frac{1}{T}\sum_{t=1}^{T} G^k(i_t, j_t) - z^k$. Ce joueur peut donc maintenant choisir l'état k, mais doit le payer au prix z^k. On montre que $\Gamma_T^*(z)$ a une valeur $w_T(z)$. w_T est convexe, et liée à la valeur du jeu primal $\Gamma_T(p)$ par les formules de conjugaison :

$$\begin{aligned} w_T(z) &= \max_{p \in \Delta(K)} (v_T(p) - \langle p, z\rangle), \\ v_T(p) &= \inf_{z \in \mathbb{R}^K} (w_T(z) + \langle p, z\rangle). \end{aligned}$$

Certains travaux privilégient une approche fonctionnelle *via* l'étude d'opérateurs (Rosenberg et Sorin 2001, voir aussi Sorin, 2002). Cette approche s'applique à la fois aux jeux répétés à information incomplète et aux jeux stochastiques.

Laraki (2004) étudie l'opérateur de convexification et s'intéresse à la préservation de la continuité et du caractère Lipschitz.

Par ailleurs, il est possible de généraliser le modèle standard de la partie 1 et de prouver l'existence de la valeur dans le cas où l'état n'est plus tiré aléatoirement une fois pour toutes au départ, mais évolue selon une chaîne de Markov observée uniquement par le joueur 1 (Renault, 2006).

Enfin, de Meyer et Moussa Saley (2003) se sont intéressés à l'origine des mouvements browniens dans les modèles financiers. Ils ont introduit un modèle de jeu de marché basé sur un jeu répété à manque d'information d'un seul côté, et prouvent l'apparition endogène d'un mouvement brownien.

5.3. Somme non nulle. — Dans le cadre de la partie 4.2 des jeux à manque d'information d'un seul côté et somme non nulle, on peut étudier le nombre d'étapes de communication nécessaires à la réalisation d'équilibres, lié à la convergence des bimartingales (Aumann et Maschler 1995, Aumann et Hart 1986, Forges 1984, Forges 1990). Indiquons que F. Forges (1988) a aussi donné une caractérisation des paiements d'équilibres, pour une notion plus générale d'équilibre appelée équilibre en communication.

Par ailleurs, on peut étudier le sous cas où chaque joueur connaît ses propres paiements. Lorsqu'il y a manque d'information d'un seul côté, cela correspond à supposer que la matrice des paiements du joueur 2 est indépendante de k. On montre (Shalev, 1994) que tout paiement d'équilibre s'obtient alors comme paiement d'équilibre complètement révélateur. Ce résultat peut se généraliser au cas de manque d'information des deux côtés et somme non nulle (voir l'article non publié de Koren, 1992), et il peut ne pas exister d'équilibre même quand les deux joueurs connaissent leurs paiements.

Un autre modèle traite du cas d'information dit symétrique. Les deux joueurs ont alors une information incomplète, mais identique, sur l'état de la nature. Ils reçoivent après chaque étape le même signal, dépendant notamment de cet état. A. Neyman et S. Sorin (1998) ont montré l'existence de *paiements* d'équilibres dans le cas de deux joueurs (pour la somme nulle, voir Forges, 1982).

Très peu d'études ont concerné le cas d'au moins 3 joueurs. On trouve un résultat partiel (pour deux états de la nature) d'existence d'équilibre de type plan joint dans Renault (2001a). Enfin, pour des modèles de jeux répétés à n joueurs à information incomplète et avec signaux, on trouve des résultats d'existence d'équilibres particuliers (complètement révélateurs) chez Renault et Tomala, 2004b (voir aussi Renault, 2001b), où la transmission stratégique d'information est étudiée indépendamment des paiements.

Bibliographie

Aumann (R.J.) & Hart (S.)
[1986] Bi-convexity and bi-martingales, *Israel Journal of Mathematics*, 54 (1986), p. 159–180.

Aumann (R.J.) & Maschler (M.)
[1995] *Repeated games with incomplete information*, M.I.T. Press, 1995 ; avec la collaboration de R. Stearns (contient une réédition de travaux de 1966,67,68).

Blackwell (D.)
[1956] An analog of the minmax theorem for vector payoffs, *Pacific Journal of Mathematics*, 65 (1956), p. 1–8.

de Meyer (B.)
[1996a] Repeated games and partial differential equations, *Mathematics of Operations Research*, 21 (1996), p. 209–236.
[1996b] Repeated games, duality and the central limit theorem, *Mathematics of Operations Research*, 21 (1996), p. 237–251.
[1998] The maximal variation of a bounded martingale and the central limit theorem, *Annales de l'Institut Henri Poincaré, Probabilités et statistiques*, 34 (1998), p. 49–59.

de Meyer (B.) & Moussa Saley (H.)
[2003] On the strategic origin of Brownian motion in finance, *International Journal of Game Theory*, 31 (2003), p. 285–319.

de Meyer (B.) & Rosenberg (D.)
[1999] « *Cavu* » and the dual game, *Mathematics of Operations Research*, 24 (1999), p. 619–626.

Forges (F.)
[1982] Infinitely repeated games of incomplete information : symmetric case with random signals, *International Journal of Game Theory*, 11 (1982), p. 203–213.
[1984] A note on Nash equilibria in repeated games with incomplete information, *International Journal of Game Theory*, 13 (1984), p. 179–187.
[1988] Communication equilibria in repeated games with incomplete information, *Mathematics of Operations Research*, 13 (1988), p. 191–231.
[1990] Equilibria with communication in a job market example, *Quarterly Journal of Economics*, 105 (1990), p. 375–398.
[1992] Repeated Games of Incomplete Information : Non-zero sum, dans Aumann (R.J.) & Hart (S.), éd., *Handbook of Game Theory, I*, Elsevier Science Publishers, 1992, p. 155–177.

Harsanyi (J.)
[1967-68] Games with incomplete information played by 'Bayesian'

players, parts I-III, *Management Science*, 8 (1967-68), p. 159–182, 320–334, 486–502.

HART (S.)

[1985] Nonzero-sum two-person repeated games with incomplete information, *Mathematics of Operations Research*, 10 (1985), p. 117–153.

KOHLBERG (E.)

[1975] Optimal strategies in repeated games with incomplete information, *International Journal of Game Theory*, 4 (1975), p. 7–24.

KOREN (G.)

[avril 1992] Two-person repeated games where players know their own payoffs, avril 1992 ; document de travail basé sur une « master thesis » à l'Université de Tel-Aviv, 50 pages, `http://www.ma.huji.ac.il/hart/papers/koren.pdf`.

LARAKI (R.)

[2001a] Variational inequalities, system of functional equations and incomplete information repeated games, *SIAM Journal on control and optimization*, 40 (2001), p. 516–524.

[2001b] The splitting game and applications, *International Journal of Game Theory*, 30 (2001), p. 359–376.

[2002] Repeated games with lack of information on one side : the dual differential approach, *Mathematics of Operations Research*, 27 (2002), p. 419–440.

[2004] On the regularity of the convexification operator on a compact set, *Journal of Convex Analysis*, 11 (2004), p. 209–234.

MAYBERRY (J.-P.)

[1967] Discounted repeated games with incomplete information, dans *Report of the U.S. Arms control and disarmament agency*, vol. ST116, chapter V, Princeton : Mathematica, 1967, p. 435–461.

MERTENS (J.-F.)

[1972] The value of two-person zero-sum repeated games : the extensive case, *International Journal of Game Theory*, 1 (1972), p. 217–227.

MERTENS (J.-F.), SORIN (S.) & ZAMIR (S.)

[1994] Repeated games, dans *CORE discussion paper*, 1994, p. 9420–9422.

MERTENS (J.-F.) & ZAMIR (S.)

[1971] The value of two-person zero-sum repeated games with lack of information on both sides, *International Journal of Game Theory*, 1 (1971), p. 39–64.

[1976] The normal distribution and repeated games, *International Journal of Game Theory*, 5 (1976), p. 187–197.

[1977] A duality theorem on a pair of simultaneous functional equations, *Journal of Mathematical Analysis and Applications*, 60 (1977), p. 550–558.

[1985] Formulation of Bayesian analysis for games with incomplete information, *International Journal of Game Theory*, 14 (1985), p. 1–29.

Neyman (A.) & Sorin (S.)

[1998] Equilibria in Repeated Games with Incomplete Information : The General Symmetric Case, *International Journal of Game Theory*, 27 (1998), p. 201–210.

Renault (J.)

[2000] 2-player repeated games with lack of information on one side and state independent signalling, *Mathematics of Operations Research*, 4 (2000), p. 552–572.

[2001a] 3-player repeated games with lack of information on one side, *International Journal of Game Theory*, 30 (2001), p. 221–246.

[2001b] Learning sets in state dependent signalling game forms : a characterization, *Mathematics of Operations Research*, 26 (2001), p. 832–850.

[2006] The value of Markov chain games with lack of information on one side, *Mathematics of Operations Research*, 31 (2006), p. 490–512.

Renault (J.) & Tomala (T.)

[2004] Learning the state of nature in repeated games with incomplete information and signals, *Games and Economic Behavior*, 47 (2004), p. 124–156.

Rosenberg (D.)

[1998] Duality and Markovian strategies, *International Journal of Game Theory*, 27 (1998), p. 577–597.

Rosenberg (D.) & Sorin (S.)

[2001] An operator approach to zero- sum repeated games, *Israel Journal of Mathematics*, 121 (2001), p. 221–246.

Shalev (J.)

[1994] Nonzero-Sum Two-Person Repeated Games with Incomplete Information and Known-Own Payoffs, *Games and Economic Behavior*, 7 (1994), p. 246–259.

Simon (R.S.)

[2002] Separation of joint plan equilibrium payoffs from the min-max functions, *Games and Economic Behavior*, 1 (2002), p. 79–102.

Simon (R.S.), Spież (S.) & Toruńczyk (H.)

[1995] The existence of equilibria in certain games, separation for families of convex functions and a theorem of Borsuk-Ulam type, *Israel Journal of Mathematics*, 92 (1995), p. 1–21.

[2002] Equilibrium existence and topology in some repeated games with incomplete information, *Transactions of the AMS*, 354 (2002), p. 5005–5026.

Sorin (S.)
[1983] Some results on the existence of Nash equilibria for non- zero sum games with incomplete information, *International Journal of Game Theory*, 12 (1983), p. 193–205.
[1984] On a pair of simultaneous functional equations, *Journal of Mathematical Analysis and Applications*, 98 (1984), p. 296–303.

Sorin (S.) & Zamir (S.)
[1985] A 2-person game with lack of information on 1 and 1/2 sides, *Mathematics of Operations Research*, 10 (1985), p. 17–23.

Sorin (S.)
[2002] *A first course on zero-sum repeated games*, Mathématiques et Applications, Springer, 2002.

Spinat (X.)
[2002] A necessary and sufficient condition for approchability, *Mathematics of Operations Research*, 27 (2002), p. 31–44.

Vieille (N.)
[1992] Weak approachability, *Mathematics of Operations Research*, 17 (1992), p. 781–791.

Zamir (S.)
[1971] On the relation between finitely and infinitely repeated games with incomplete information, *International Journal of Game Theory*, 1 (1971), p. 179–198.
[1973] On repeated games with general information function, *International Journal of Game Theory*, 21 (1973), p. 215–229.
[1992] Repeated Games of Incomplete Information : zero-sum, dans Aumann (R.J.) & Hart (S.), éd., *Handbook of Game Theory, I*, Elsevier Science Publishers, 1992, p. 109–154.

J. Renault, Ceremade, Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris cedex 16
E-mail : `renault@ceremade.dauphine.fr`
Url : `http://www.ceremade.dauphine.fr/~renault/`