

# Étude d'un jeu de poursuite plane avec connaissance imparfaite d'une coordonnée

On a planar pursuit game  
with imperfect knowledge of a coordinate

Alain Rapaport et Pierre Bernhard

INRIA Sophia-Antipolis

2004, route des Lucioles - BP 93

06902 Sophia-Antipolis - France

---

*RÉSUMÉ* : Un exemple simple de poursuite dans le plan, où la solution en information parfaite est déjà connue, est étudié dans le cadre d'une structure d'information incomplète. La théorie non linéaire, proposée par Bernhard pour les jeux en information imparfaite, est utilisée ici dans une version à temps final variable.

*ABSTRACT* : We investigate a simple 2-D pursuit game whose full state feed-back solution is elementary. We consider a partial information case, extending to the variable end-time case our earlier theory for fixed end-time partial information games.

---

## 1 Le jeu considéré

### 1.1 Le jeu en information parfaite

Considérons le très classique problème de poursuite-évasion dans le plan [1] où poursuivant et fugitif se déplacent à vitesse constante, respectivement  $v_P$  et  $v_F$ , et contrôlent l'angle de leur vecteur vitesse, respectivement  $u \in U = [0, 2\pi]$  et  $v \in V = [0, 2\pi]$ . Dans un repère centré sur le poursuivant, les équations dynamiques s'écrivent :

$$S(x_0, y_0, u(\cdot), v(\cdot)) : \begin{cases} x(0) &= x_0 \\ y(0) &= y_0 \\ \dot{x}(t) &= v_F \cos(v(t)) - v_P \cos(u(t)) \\ \dot{y}(t) &= v_F \sin(v(t)) - v_P \sin(u(t)) \end{cases}$$

La capture a lieu lorsque les positions des deux joueurs coïncident :  $x = y = 0$ . Le critère que les poursuivant et fugitif cherchent respectivement à minimiser et maximiser est le temps de capture :

$$t_c(x_0, y_0, u(\cdot), v(\cdot)) = \inf_{t \geq 0} \{t \mid x(t) = 0 \text{ et } y(t) = 0\}$$

lorsque  $(x(\cdot), y(\cdot))$  est solution de  $S(x_0, y_0, u(\cdot), v(\cdot))$ .

Si  $\mathcal{U}$  et  $\mathcal{V}$  désignent les ensembles de stratégies en boucle ouverte, i.e. les fonctions mesurables du temps  $u : \mathbb{R} \rightarrow U$  et  $v : \mathbb{R} \rightarrow V$ , les classes de stratégies envisagées sont les retours d'état admissibles :  $\Phi = \{\phi : X \rightarrow U\}$  et  $\Psi = \{\psi : X \rightarrow V\}$  tels que  $\forall (x_0, y_0) \in X$ , les deux systèmes suivants :

$$\begin{aligned} S(x_0, y_0, u(\cdot), \psi(\cdot)) , \forall u \in \mathcal{U} , \forall \psi \in \Psi \\ S(x_0, y_0, \phi(\cdot), v(\cdot)) , \forall \phi \in \Phi , \forall v \in \mathcal{V} \end{aligned}$$

admettent une unique trajectoire solution ( $X = \mathbb{R}^2$  est l'espace d'état).

Lorsque  $v_P > v_F$ , la théorie d'Isaacs nous enseigne qu'il existe un point selle :

$$t_c^*(x_0, y_0) = \min_{\phi \in \Phi} \max_{\psi \in \Psi} t_c(x_0, y_0, \phi(\cdot), \psi(\cdot)) = \max_{\psi \in \Psi} \min_{\phi \in \Phi} t_c(x_0, y_0, \phi(\cdot), \psi(\cdot))$$

qui garantit au poursuivant une capture en temps fini.

En effet, l'équation d'Isaacs s'écrit :

$$\begin{cases} 1 - (v_P - v_F) \sqrt{\left(\frac{\partial t_c^*}{\partial x}\right)^2 + \left(\frac{\partial t_c^*}{\partial y}\right)^2} = 0 \\ t_c^*(0, 0) = 0 \end{cases}$$

dont on vérifie aisément que :

$$t_c^*(x, y) = \frac{\sqrt{x^2 + y^2}}{v_P - v_F}$$

est solution. Dès lors, les contrôles optimaux  $(\phi^*, \psi^*)$  sont :

$$\begin{cases} \cos(\phi^*(x, y)) = \cos(\psi^*(x, y)) = \frac{x}{\sqrt{x^2 + y^2}} \\ \sin(\phi^*(x, y)) = \sin(\psi^*(x, y)) = \frac{y}{\sqrt{x^2 + y^2}} \end{cases}$$

Ainsi, les stratégies optimales consistent pour les deux joueurs à se déplacer en ligne droite dans la direction  $\vec{PF}$ .

## 1.2 Le jeu en information imparfaite

Dans le même contexte ( $v_P > v_F$ ), considérons le jeu où le poursuivant ne possède qu'une connaissance imparfaite de l'état courant :  $y = h(x, y)$ . Nous supposons de plus qu'une information a priori sur l'état initial lui est fournie :

$$x_0 \in X_0 = [m_0 - l_0, m_0 + l_0]$$

Nous nous plaçons dans le cadre de la théorie développée par Bernhard [2], (et simultanément par James et al. [5]) c'est-à-dire que nous envisageons les classes de contrôleurs suivantes :

- $v$  joue en boucle ouverte :  $v(\cdot) \in \mathcal{V}$ . Pour une étude dans le pire des cas, il est commode de considérer  $x_0$  comme relevant également du choix de l'adversaire.  $\Omega(m_0, l_0)$  désigne l'ensemble des paires  $\omega = (x_0, v(\cdot)) \in X_0 \times \mathcal{V}$ , que l'on nomme *perturbations* en hommage à la littérature sur le contrôle robuste (dans le mesure où on s'intéresse ici au minimax, et non au point selle, il n'y a aucune différence entre prendre  $v(\cdot) \in \mathcal{V}$  ou autoriser des stratégies de rétroaction).
- $u$  joue en boucle fermée sur  $y$ .  $\mathcal{M}$  est l'ensemble des contrôleurs causaux :  $u(t) = \mu_t(y^t)$  où  $y^t = \{y(\tau) | \tau \leq t\}$  est l'historique des observations, tels que :  $\forall x_0 \in X_0, \forall v(\cdot) \in \mathcal{V}$ , le système  $S(x_0, y_0, \mu(\cdot), v(\cdot))$  possède une unique trajectoire solution.

Le jeu se termine lorsqu'il y a capture, information que le joueur  $u$  ne connaît pas a priori. Néanmoins, comme on s'intéresse à un contrôleur dans le pire des cas, la capture ne peut se produire que plus tôt si son opposant ne joue pas au mieux, et les actions ultérieures de  $u$  deviennent alors sans objet.

Nous nous intéressons à l'existence d'un contrôleur min-max :

$$\bar{t}_c(m_0, l_0, y_0) = \min_{\mu \in \mathcal{M}} \max_{\substack{(x_0, v) \in \\ \Omega(m_0, l_0)}} t_c(x_0, y_0, \mu(\cdot), v(\cdot))$$

En suivant une idée initiée simultanément par James et Baras [5], pour un critère de la forme :

$$J(t_0, x_0, y_0, u(\cdot), v(\cdot)) = M(x(t_c), y(t_c), t_c) + \int_{t_0}^{t_c} L(\tau, x(\tau), y(\tau), u(\tau), v(\tau)) d\tau + N(x_0)$$

l'information avec laquelle le joueur  $u$  bâtit une boucle fermée optimale est la fonction *pire dépense conditionnelle* :

$$W_t(\xi) = \max_{\substack{(x_0, v) \in \\ \Omega_t(\xi)}} \left\{ \int_{t_0}^t L(\tau, x(\tau), y(\tau), u(\tau), v(\tau)) d\tau + N(x_0) \right\}$$

où  $\Omega_t(\xi)$  est l'ensemble des perturbations compatibles avec les observations passées  $y^t$  et qui conduisent le système dans l'état  $\xi$  à l'instant  $t$ . Pour un problème en temps minimal, en choisissant  $M(\cdot)$ ,  $L(\cdot)$  et  $N(\cdot)$  comme suit :

$$\begin{cases} M(x, y, t) & = t \\ L(t, x, y, u, v) & = 0 \\ N(x_0) & = \begin{cases} 0 & \text{si } x_0 \in X_0 \\ -\infty & \text{sinon} \end{cases} \end{cases}$$

la fonction  $W_t(\cdot)$  est une application caractéristique d'un certain ensemble  $X_t$ . En effet, fixons un contrôleur  $\mu \in \mathcal{M}$  et une perturbation  $\omega \in \Omega$ . Joués ensemble jusqu'à l'instant  $t$ , ils génèrent des observations  $y^t = \{y(\tau) | \tau \leq t\}$  et une synthèse du contrôle  $u^t = \{u(\tau) | \tau \leq t\}$ . Puisque les perturbations passées compatibles avec les observations, qui conduiraient à la capture exactement à l'instant  $t$ , ne sont plus valides dans le

futur (si le jeu continue, c'est qu'il n'y a pas eu capture),  $\Omega_t[u^t, y^t]$  est l'ensemble des perturbations compatibles avec les observations et n'ayant jamais rencontré la cible :

$$\Omega_t[u^t, y^t] = \{(x_0, v(\cdot)) \in X_0 \times \mathcal{V} \mid \forall \tau \leq t, \dot{y}(\tau) = v_F \sin(v(\tau)) - v_P \sin(u(\tau)) \text{ et } (x(\tau), y(\tau)) \neq (0, 0)\}$$

Si  $\Omega_t[u^t, y^t](\xi)$  désigne l'ensemble conditionné à amener le système dans l'état  $\xi$  à l'instant  $t$  :

$$\Omega_t[u^t, y^t](\xi) = \{\omega \in \Omega_t[u^t, y^t] \mid x(t) = \xi\}$$

son support :

$$X_t[u^t, y^t] = \{\xi \mid \Omega_t[u^t, y^t](\xi) \neq \emptyset\}$$

représente alors l'ensemble des états que le système peut prendre à l'instant  $t$ , conditionnellement aux observations effectuées, et la fonction pire dépense conditionnelle devient :

$$W_t(x) = \begin{cases} 0 & \text{si } x \in X_t \\ -\infty & \text{sinon.} \end{cases}$$

La cible à atteindre est alors :

$$\mathbb{R}^+ \times \{W^\emptyset\} \quad (\text{où } W^\emptyset(x) = -\infty, \forall x \in X),$$

et le critère :

$$t_c = \inf_{t \geq 0} \{t \mid X_t = \emptyset\}.$$

Notre propos est d'écrire une équation de type Hamilton-Jacobi-Isaacs, comme dans [2] où  $W \in \mathcal{W}$  (ensemble des fonctions  $W : X \rightarrow \mathbb{R} \cup \{-\infty\}$  atteignables lorsque  $\mu, \omega$  et  $t$  parcourent respectivement  $\mathcal{M}, \Omega$  et  $\mathbb{R}^+$ ) est l'état du système, et dont  $U : \mathbb{R}^+ \times \mathcal{W} \rightarrow \mathbb{R}^+$ , fonction valeur du jeu, serait solution.

Pour étendre les résultats de Bernhard [2] à ce cadre précis, une difficulté majeure se présente : l'écriture du *filtre d'évolution* de la fonction *pire dépense conditionnelle* :

$$\forall (t, \xi), \quad \frac{\partial W_t}{\partial t}(\xi) = F_t[W_t, u, y](\xi).$$

En effet, on ne peut imposer ici que les fonctions  $W : (\xi, t) \in X \times \mathbb{R} \mapsto W_t(\xi)$  soient  $C^1$ , puisque cette hypothèse leur interdirait de prendre la valeur  $-\infty$  (tous les états  $\xi \in X$  seraient alors accessibles). Nous proposons néanmoins une équation d'Isaacs pour des problèmes à temps final variable, sous des hypothèses de régularité de la fonction  $W$  et du filtre d'évolution, puis nous montrons sur l'exemple comment une représentation en dimension finie des ensembles  $X_t$  permet finalement d'écrire une telle équation, dont on donne enfin la solution.

## 2 L'équation d'Isaacs pour le jeu en temps final variable

### 2.1 Le problème envisagé

Soit un jeu différentiel  $\mathcal{J}(t_0, X_0)$  :

$$\begin{aligned} x_0 &\in X_0 \subset X \\ \dot{x} &= f(t, x, u, v), \quad x \in X = \mathbb{R}^n, \quad u \in U \subset \mathbb{R}^p, \quad v \in V \subset \mathbb{R}^q, \quad t \geq t_0 \\ J(t_0, x_0, u(\cdot), v(\cdot)) &= M(t_c, x(t_c)) + \int_{t_0}^{t_c} L(\tau, x(\tau), u(\tau), v(\tau)) d\tau + N(x_0) \\ t_c(t_0, x_0, u, v) &= \inf_{t \geq t_0} \{ t \mid (t, x(t)) \in \mathcal{C} \} \end{aligned}$$

(où  $U$  et  $V$  sont des compacts, et  $\mathcal{C} \subset \mathbb{R} \times X$  est la cible)  
et une observation :

$$\mathcal{O} : \begin{cases} y(t) = h(t, x(t), v(t)) \in Y \subset \mathbb{R}^p \quad r \leq n \\ y^t = \{ y(\tau) \mid \tau \leq t \}. \end{cases}$$

( $h$  est supposée être surjective :  $Y = \text{Im}(h)$ ). On notera  $S(t_0, x_0, u, v)$  l'ensemble des trajectoires solutions du système dynamique.

On s'intéresse aux classes de contrôleurs :

$$\begin{cases} \mathcal{M} = \{ \mu : (t, y^t) \in \mathbb{R} \times \mathcal{Y}^t \mapsto U \} \\ \Omega = X_0 \times \mathcal{V} \end{cases}$$

telles que pour toute paire  $(\mu, \omega)$  le système différentiel admette une unique solution ( $\mathcal{V}$  et  $\mathcal{Y}$  désignent les ensembles des boucles ouvertes et observations possibles). Dans la suite, nous procéderons à l'abus de notation :  $(t_0, x_0, u, v) = (t_0, u, \omega)$ .

On cherche à caractériser la fonction valeur :

$$U(t_0, N) = \min_{\mu \in \mathcal{M}} \max_{\omega \in \Omega} J(t_0, \mu, \omega).$$

### 2.2 Quelques définitions

**Définition 1 : Ensemble des perturbations compatibles avec les observations :**

$$\Omega_t[u^t, y^t] = \{ \omega \in \Omega \mid \forall \tau \leq t, y(\tau) = h(\tau, x(\tau), v(\tau)) \quad , \quad x(\cdot) \in S(t_0, u, \omega) \}$$

**Remarque :** Nous n'imposons pas ici qu'une perturbation  $\omega \in \Omega_t$  ait évité la capture avant l'instant  $t$ , afin d'obtenir un peu plus de régularité de la fonctionnelle  $W_t$ . Les conditions de capture seront envisagées plus loin en termes de  $W_t$ .

**Définition 2 : Ensemble des perturbations conditionnelles :**

$$\forall (t, \xi) \in \mathbb{R} \times X, \quad \Omega_t[u^t, y^t](\xi) = \{ \omega \in \Omega_t[u^t, y^t] \mid x(t) = \xi \}$$

où  $x(\cdot) \in S(t_0, u^t, \omega)$ .

**Définition 3 : Fonction “Pire dépense conditionnelle” :**

Pour une paire  $(u^t, y^t) \in \mathcal{U}^t \times \mathcal{Y}^t$  fixée,

$$\forall \xi \in X, \quad W_t(\xi) = \sup_{\omega \in \Omega_t(\xi)} \left\{ \int_{t_0}^t L(\tau, x(\tau), u(\tau), v(\tau)) d\tau + N(x_0) \right\}$$

avec  $\omega = (x_0, v)$ ,  $x(\cdot) \in S(t_0, x_0, u, v)$  et la convention suivante :

$$\Omega_t(\xi) = \emptyset \implies W_t(\xi) = -\infty.$$

**Définition 4 : Ensemble des états compatibles avec les observations :**

$$\forall t \geq t_0, \quad X_t[u^t, y^t] = \{ \xi \in X \mid \Omega_t[u^t, y^t](\xi) \neq \emptyset \},$$

**Définition 5 : Ensemble d’atteignabilité compatible avec les observations :**

$$\mathcal{R}[u^T, y^T] = \bigcup_{t \in [t_0, +\infty[} \{ \{t\} \times X_t[u^t, y^t] \}.$$

Afin d’étudier le problème en termes de pires dépenses conditionnelles, sont définis les ensembles fonctionnels suivants :

**Définition 6 : Ensembles des pires dépenses :**

$$\forall t \geq t_0, \quad \mathcal{W}_t = \{ W_t[u^t, y^t] \mid u^t \in \mathcal{U}^t, \omega^t \in \Omega^t \},$$

**Définition 7 : Ensemble d’atteignabilité des pires dépenses :**

$$\mathcal{Z} = \bigcup_{t \in [t_0, +\infty[} \{ \{t\} \times \mathcal{W}_t \}.$$

**2.3 Quelques hypothèses**

**Hypothèse 1 ( Forte capturabilité ) :**  $\forall (t_0, X_0) \in \mathbb{R} \times \mathcal{P}(X)$ ,

$$\exists T(t_0, X_0) < +\infty, \exists u \in \mathcal{U}, \forall x_0 \in X_0, \forall v \in \mathcal{V}, t_c(t_0, x_0, u, v) < T(t_0, X_0).$$

Fixons une paire  $(t_0, X_0)$  et posons  $T = T(t_0, X_0)$ ,

**Hypothèse 2 (Différentiabilité de la fonction  $W_t$ ) :**

**a)** Pour tout  $t \in [t_0, T]$  et  $(u^t, y^t)$ , l’application

$$W_t(\cdot) : X_t[u^t, y^t] \longrightarrow \mathbb{R} \cup \{-\infty\}$$

est bien définie, et le sup est atteint :

$$\forall \xi \in X_t[u^t, y^t], \exists \widehat{\omega} = (\widehat{v}, \widehat{x}_0) \in \Omega_t[u^t, y^t](\xi) :$$

$$W_t(\xi) = \int_{t_0}^t L(\tau, x(\tau), u(\tau), \widehat{v}(\tau)) d\tau + N(\widehat{x}_0)$$

b) A  $u^T$  et  $y^T$  fixés, pour tout  $t \in [t_0, T]$  et tout  $\xi \in X_t[u^t, y^t]$ , l'application  $W_t(\xi)$  est différentiable en  $t$  :

$$\forall (t, \xi) \in \mathcal{R}[u^T, y^T], \quad \frac{\partial W_t(\xi)}{\partial t} = F_t[W_t, u_t, y_t](\xi).$$

Afin d'exhiber une des pires perturbations pour le problème dont la fonction valeur est une application  $U : \mathcal{Z} \rightarrow \mathbb{R}$ , nous avons besoin de l'hypothèse d'existence suivante :

**Hypothèse 3** Pour tout contrôleur  $\mu \in \mathcal{M}$ , et toute fonction  $U : \mathcal{Z} \rightarrow \mathbb{R}$  qui admet, pour tout  $(t, W) \in \mathcal{Z}$  une dérivée de Gâteaux  $D_W U(t, W)$  par rapport à  $W$ , on suppose qu'il existe une perturbation  $\bar{\omega} \in \Omega$  telle que :

$$\forall t, \quad \bar{y}(t) \in \arg \max_{y \in Y} D_W U(t, \bar{W}_t) F_t[\bar{W}_t, \bar{u}(t), y]$$

où  $(\bar{u}^t, \bar{y}^t)$  est une synthèse de la paire  $(\mu, \bar{\omega})$ , et  $W_t$  la fonction pire dépense conditionnelle associée.

**Remarque :** Cette hypothèse traduit l'existence d'une perturbation  $\omega$  générant une trajectoire  $x(\cdot)$  admissible telle que l'observation  $y(\cdot)$  produite vérifie la propriété voulue, ceci pour un contrôleur  $\mu$  et une fonction  $U$  fixés. Si pour tout  $(t, x) \in \mathbb{R} \times X$ , l'application  $v \mapsto h(t, x, v) \in Y$  est surjective, alors on choisit  $v^*(t)$  de façon causale dans  $h^{-1}(t, x, \arg \max_{y \in Y} (D_W U(t, \bar{W}_t) F_t[\bar{W}_t, \bar{u}(t), y]))$ . Sinon, sous l'hypothèse que le contrôleur  $\mu$  accède à la connaissance de  $\dot{y}$ , on peut espérer que l'application  $v \mapsto h_x(t, x, v) f(t, x, u, v)$  soit surjective en  $v$ , et à défaut étudier les dérivées successives de la sortie  $y$ .

Une difficulté apportée par le temps final variable est l'écriture de la "cible" en termes de  $(t, W)$ . Nous proposons deux variantes de l'équation d'Isaacs, selon la nature de l'information sur la capture disponible.

## 2.4 Première version

Sous l'hypothèse :

**Hypothèse 4 (Information de la capture) :**

$$\forall (T, W_T) \in \mathcal{Z}, \exists t_c(W_T) \mid \forall (u^T, y^T), \forall \omega \in \Omega_T[u^T, y^T], \\ t_c(t_0, u, \omega) = t_c(W_T[u^T, y^T]),$$

Un exemple typique où cette hypothèse est validée est le cas où la cible est définie uniquement en termes de la variable  $y$  :

$$\exists \mathcal{C}_Y \subset \mathbb{R} \times Y \quad \mid \quad \mathcal{C} = \{(t, \xi) \in \mathbb{R} \times X \mid \forall v \in V, (t, h(t, \xi, v)) \in \mathcal{C}_Y\},$$

on peut définir une cible en termes de pires dépenses conditionnelles :

**Définition 8 : Ensemble cible de pires dépenses :**

$$\mathcal{C}_W = \{ (t, W_t) \in \mathcal{Z} \mid \exists (u^T, y^T) \in \Pi \times \mathcal{Y}, W_t = W_t[u^T, y^T], \exists \omega \in \Omega_T[u^T, y^T] \text{ et } t = t_c(t_0, u, \omega) \}.$$

Nous sommes maintenant en mesure d'énoncer l'équation fondamentale de la *programmation dynamique* comme dans [2], mais adaptée ici au caractère non nécessairement  $C^1$  des fonctions  $W$ .

**Proposition 1** *Sous les hypothèses 1, 2 et 4, s'il existe une fonction  $U : \mathcal{Z} \rightarrow \mathbb{R}$  vérifiant l'hypothèse 3 telle que :*

$$\begin{cases} \forall (t, W_t) \in \mathcal{Z} \setminus \mathcal{C}_W, & \frac{\partial U(t, W_t)}{\partial t} + \min_{u \in U} \max_{y \in Y} D_W U(t, W_t) F_t[W_t, u, y] = 0 \\ \forall (t, W_t) \in \mathcal{C}_W, & U(t, W_t) = \max_{\xi \in X} \{M(t, \xi) + W_t(\xi)\} \end{cases} \quad (1)$$

et si  $\hat{\mu}_t(W_t) \in \arg \min_u$  est un contrôleur admissible, alors  $\hat{\mu}$  est un contrôleur optimal, garantissant un coût  $U(t_0, N)$ . De plus,  $t_c = \arg \inf_{t \geq t_0} \{(t, W_t) \in \mathcal{C}_W\}$  est le temps de capture correspondant.

**Preuve**

- Remarquons tout d'abord que le contrôleur proposé  $\hat{\mu}^t$  est bien causal, puisqu'il ne dépend que de  $W_t$ .

- Supposons que  $u$  opte pour la stratégie  $\hat{\mu}_t(\cdot)$  et soit une perturbation  $\bar{\omega} = (\bar{x}_0, \bar{v}(\cdot)) \in \Omega$ , alors la paire  $(\hat{\mu}, \bar{\omega})$  génère une trajectoire  $\{x_t\}$ , un temps de capture  $t_c$ , des observations  $\{y_t\}$ , et des fonctions pire dépense conditionnelle  $\{W_t\}$  :

Grâce à l'équation 1, on a :

$$\frac{\partial U(t, W_t)}{\partial t} + D_W U(t, W_t) F_t[W_t, \hat{\mu}_t(W_t), y_t] \leq 0.$$

Donc l'application  $t \mapsto U(t, W_t)$  est décroissante (au sens large), donc  $U(t_c, W_{t_c}) \leq U(t_0, W_{t_0})$ .

D'autre part,

$$\begin{aligned} U(t_c, W_{t_c}) &= \max_{\xi \in X} [M(t_c, \xi) + W_{t_c}(\xi)] \\ &= \max_{\xi \in X} \left[ M(t_c, \xi) + \max_{\substack{(x_0, v) \in \\ \Omega_{t_c}[\hat{\mu}, \bar{\omega}]}(\xi)}} \left\{ \int_{t_0}^{t_c} L(\tau, x(\tau), u(\tau), v(\tau)) d\tau + N(x_0) \right\} \right] \\ &= \max_{\omega \in \Omega_{t_c}[\hat{\mu}, \bar{\omega}]} \left[ M(t_c, \xi) + \int_{t_0}^{t_c} L(\tau, x(\tau), u(\tau), v(\tau)) d\tau + N(x_0) \right] \end{aligned}$$



Or, par construction même du processus  $\Omega_{(\cdot)}$ , on a  $:\bar{\omega} \in \Omega_{t_c}[\hat{\mu}, \bar{\omega}]$ . Donc :

$$U(t_c, W_{t_c}) \geq J(t_0, \hat{\mu}, \bar{\omega})$$

et par conséquent :

$$U(t_0, N) \geq J(t_0, \hat{\mu}, \bar{\omega}).$$

• Soit un contrôleur quelconque  $\mu_t(\cdot)$ , et une pire perturbation vérifiant l'hypothèse 3. Soient  $\bar{u}$  et  $\hat{y}$  les boucles ouvertes alors engendrées, et  $\widehat{W}_t[\bar{u}^t, \hat{y}^t]$  la pire dépense conditionnelle associée. Pour cette fonction  $\widehat{W}_t$ , l'instant de capture  $t_c$  est défini de façon univoque, grâce à l'hypothèse 4.

L'équation 1 nous enseigne :

$$\forall t \leq t_c, \quad \frac{\partial U(t, \widehat{W}_t)}{\partial t} + D_W U(t, \widehat{W}_t) F_t[\widehat{W}_t, \bar{u}_t(\widehat{W}_t), \hat{y}_t] \geq 0.$$

Ainsi, l'application  $t \mapsto U(t, \widehat{W}_t)$  est croissante, et par conséquent :

$$U(t_c, \widehat{W}_{t_c}[\bar{u}, \hat{y}]) \geq U(t_0, N).$$

Puisque  $\forall t \leq t_c, X_t[\bar{u}^{t_c}, \hat{y}^{t_c}] \neq \emptyset$ , on peut définir pour tout  $\xi \in X_t[\bar{u}^{t_c}, \hat{y}^{t_c}]$  :

$$\omega_t^*(\widehat{W}_t, \xi, \bar{u}^t, \hat{y}^t) \in \arg \max_{\omega \in \Omega_t[\bar{u}^t, \hat{y}^t](\xi)} \left\{ \int_{t_0}^t L(\tau, x(\tau), u(\tau), v(\tau)) d\tau + N(x_0) \right\}.$$

Ainsi, pour  $\xi \in X_{t_c}[\bar{u}^{t_c}, \hat{y}^{t_c}]$ , on a :

$$J(t_0, \mu, \omega_{t_c}^*(\widehat{W}_{t_c}, \xi, \bar{u}^{t_c}, \hat{y}^{t_c})) = M(t_c, \xi) + \widehat{W}_{t_c}(\xi).$$

Soit

$$\hat{\xi} \in \arg \max_{\xi \in X_{t_c}[\bar{u}^{t_c}, \hat{y}^{t_c}]} \left[ M(t_c, \xi) + \widehat{W}_{t_c}(\xi) \right],$$

la perturbation correspondante  $\hat{\omega} = \omega_{t_c}^*(\widehat{W}_{t_c}, \hat{\xi}, \bar{u}^{t_c}, \hat{y}^{t_c})$  permet d'obtenir :

$$J(t_0, \mu, \hat{\omega}) = \max_{\xi \in X_{t_c}[\bar{u}^{t_c}, \hat{y}^{t_c}]} \left[ M(t_c, \xi) + \widehat{W}_{t_c}(\xi) \right] = U(t_c, \widehat{W}_{t_c}) \geq U(t_0, N).$$

• Le contrôleur  $\hat{\mu}$ , contre une perturbation  $\hat{\omega} = \omega_{t_c}^*(\widehat{W}_{t_c}, \hat{\xi}, \bar{u}^{t_c}, \hat{y}^{t_c})$  déterminée comme précédemment, garantit le coût :

$$J(t_0, \hat{\mu}, \hat{\omega}, \hat{y}^{t_c}) = U(t_0, N).$$

□

## 2.5 Deuxième version

Soit l'ensemble des temps de capture compatibles avec les observations générées par une paire  $(\mu, \omega)$  :

### Définition 9 (Ensemble des temps de capture) :

Soient  $(\mu, \omega) \in \mathcal{M} \times \Omega$ ,  $u^t$  et  $y^t$  désignent les synthèses en boucle ouverte du contrôle  $u$  et des observations effectuées,

$$\mathcal{T}_C[\mu, \omega] = \{ t_c \mid \exists \xi \in X_{t_c}[u^T, y^T], (t_c, \xi) \in \mathcal{C} \}.$$

Le problème peut se réécrire ainsi :

Pour tout  $x_0 \in X_0$  tel que  $(t_0, x_0) \notin \mathcal{C}$ , existe-t-il un contrôleur  $\mu^* \in \mathcal{M}_{[t_0, T]}$  garantissant :

$$U(t_0, X_0) = \min_{\mu \in \mathcal{M}} \max_{\substack{(\omega, t_c) \in \\ \Omega \times \mathcal{T}_C[\mu, \omega]}} M(t_c, x(t_c)) + \int_{t_0}^{t_c} L(\tau, x(\tau), u(\tau), v(\tau)) d\tau + N(x_0)$$

et si oui, comment le construire ?

**Proposition 2** Sous les hypothèses 1 et 2, s'il existe une fonction  $U : \mathcal{Z} \rightarrow \mathbb{R}$  vérifiant l'hypothèse 3 telle que  $\forall (t, W_t) \in \mathcal{Z}$ ,

- i)  $U(t, W_t) \geq \max_{\xi \in X} [M(t, \xi) + W_t(\xi)]$
- ii)  $\frac{\partial U(t, W_t)}{\partial t} + \min_{u \in U} \max_{y \in Y} D_W U(t, W_t) F_t[W_t, u, y] \geq 0$
- iii)  $U(t, W_t) \neq \max_{\xi \in X} [M(t, \xi) + W_t(\xi)] \implies$   
 $\frac{\partial U(t, W_t)}{\partial t} + \min_{u \in U} \max_{y \in Y} D_W U(t, W_t) F_t[W_t, u, y] = 0$
- iv)  $\forall (\mu, \omega) \in \mathcal{M} \times \Omega,$

$$\inf_{t \geq t_0} \left\{ t \mid U(t, W_t[u^t, y^t]) = \max_{\xi \in X} [M(t, \xi) + W_t[u^t, y^t](\xi)] \right\} = t_c(t_0, \mu, \omega)$$

et si  $\hat{\mu}_t(W_t) \in \arg \min_u$  est un contrôleur admissible, alors  $\hat{\mu}$  est un contrôleur optimal, garantissant un coût  $U(t_0, N)$ .

### Preuve

Elle est analogue à celle de la proposition 1 :

• Soient  $(\mu, \omega) \in \mathcal{M} \times \Omega$ , et  $(t_c, x_{t_c})$  la capture réelle. Grâce à i), on a :

$$\begin{aligned} U(t_c, W_{t_c}) &\geq \max_{\xi \in X} [M(t_c, \xi) + W_{t_c}(\xi)] \\ &\geq M(t_c, x_{t_c}) + W_{t_c}(x_{t_c}) \\ &\geq J(t_0, \mu, \omega). \end{aligned}$$

Si  $\widehat{\mu}$  est choisi par le joueur u, à l'aide de iii) et iv) on déduit :

$$U(t_c, W_{t_c}) \leq U(t_0, N)$$

d'où :

$$U(t_0, N) \geq J(t_0, \widehat{\mu}, \omega).$$

• Choisissons un contrôleur  $\mu$  et associons lui une perturbation selon l'hypothèse 3, dont découle une certaine fonction  $\widehat{W}_t[\widehat{u}^t, \widehat{y}^t]$ . Pour tout  $t_c \in \mathcal{T}_C[\widehat{u}^T, \widehat{y}^T]$ , on a grâce à ii) :

$$U(t_c, \widehat{W}_{t_c}[\widehat{u}^{t_c}, \widehat{y}^{t_c}]) \geq U(t_0, N)$$

Pour  $\xi \in X_{t_c}[\widehat{u}^{t_c}, \widehat{y}^{t_c}]$ , on définit :

$$\omega_{t_c}^*(\widehat{W}_{t_c}, \xi, \widehat{u}^{t_c}, \widehat{y}^{t_c}) \in \arg \max_{\omega \in \Omega_{t_c}[\widehat{u}^{t_c}, \widehat{y}^{t_c}](\xi)} \left\{ \int_{t_0}^{t_c} L(\tau, x(\tau), u(\tau), v(\tau)) d\tau + N(x_0) \right\}$$

A l'aide de iv), on déduit qu'il existe  $(\widehat{t}_c, \widehat{\xi}_{t_f}) \in \mathcal{T}_C[\widehat{u}^T, \widehat{y}^T] \times X_{t_f}[\widehat{u}^T, \widehat{y}^T]$  tel que

$$\max_{\xi \in X} [M(\widehat{t}_c, \xi) + \widehat{W}_{\widehat{t}_c}(\xi)] = M(\widehat{t}_c, \widehat{\xi}_{t_f}) + \widehat{W}_{\widehat{t}_c}(\widehat{\xi}_{t_f}) \text{ et } (\widehat{t}_c, \widehat{\xi}_{t_f}) \in \mathcal{C}.$$

Alors  $\widehat{\omega} = \omega_{\widehat{t}_c}^*(\widehat{W}_{\widehat{t}_c}, \widehat{\xi}_{t_f}, \widehat{u}^{\widehat{t}_c}, \widehat{y}^{\widehat{t}_c})$  est une perturbation qui garantit :

$$J(t_0, \mu, \widehat{\omega}) = U(\widehat{t}_c, \widehat{W}_{\widehat{t}_c}) \geq U(t_0, N).$$

• Lorsque  $\widehat{\mu}$  est utilisé, on choisit une perturbation  $\widehat{\omega}$  comme précédemment, mais en s'assurant que  $\widehat{t}_c = t_c[\widehat{\mu}, \widehat{\omega}]$  (ceci est possible grâce à la condition iv)). Ainsi, iii) et iv) permettent de conclure :

$$J(t_0, \widehat{\mu}, \widehat{\omega}) = U(t_0, N).$$

□

**Remarque :** La condition iv) de la proposition 2 est une forme d'hypothèse d'*information* de la capture, puisque deux perturbations indiscernables par le processus  $\{W_t\}$  doivent mener au même temps de capture.

### 3 Résolution de l'exemple

Comme  $\mathcal{Z}$  est un espace fonctionnel, le contrôleur  $\widehat{\mu}$  de la proposition 2 est a priori de dimension infinie. De plus, on ne sait pas non plus, a priori, comment écrire le filtre  $F$ . Néanmoins, pour ce problème précis, tant que  $y(\cdot)$  ne s'annule pas, l'ensemble  $X_t$  conserve la structure d'un intervalle  $[m_t - l_t, m_t + l_t] \times \{y_t\}$  dont les équations dynamiques sont les suivantes :

$$\begin{cases} \dot{m}(t) &= -v_P \cos(u(t)) \\ \dot{l}(t) &= v_F |\cos(v(t))| \\ \dot{y}(t) &= v_F \sin(v(t)) - v_P \sin(u(t)) \end{cases} = \sqrt{v_F^2 - (\dot{y}(t) + v_P \sin(u(t)))^2}$$

Lorsque  $y$  s'annule et  $0 \in X_t$ , l'ensemble  $X_t$  perd son caractère connexe, et l'écriture du filtre se complique. Néanmoins, nous allons montrer que la variété  $\{y = 0\}$  est attractive et que, par conséquent, on peut décomposer la poursuite optimale en deux sous-problèmes : rejoindre la variété  $y = 0$  puis y rester. Ainsi, sous l'hypothèse que  $u$  accède à  $\dot{y}$  (cf la remarque associée à l'hypothèse 3 lorsque  $h$  n'est pas surjectif en  $v$ ), on peut appliquer la proposition 1 avec  $\{y = 0\}$  comme cible, et  $(m, l, y)$  comme paramétrage en dimension trois de l'espace  $\mathcal{Z}$ .

**Proposition 3**  $\forall \mu \in \mathcal{M}, \forall (m_0, l_0, y_0) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}$ , pour toute pire des perturbations  $(x_0, v) \in \Omega(m_0, l_0)$  il existe  $t < t_c(x_0, y_0, \mu, v)$  tel que

$$y(t) = 0.$$

### Preuve

Commençons par éliminer le cas exceptionnel où  $X_{t_c}$  serait un singleton. Ceci n'est possible que si  $l_0 = 0$  et  $\cos(v(t)) = 0, \forall t$ .  $u$  possède alors une information parfaite de l'état :  $x(t) = m(t)$ . Or, la résolution du jeu en information parfaite nous enseigne qu'un tel contrôle  $v$  n'est optimal que si  $u$  joue  $u^*$  et que si  $x_0 = m_0 = 0$ . Le coût associé est alors  $t_c^*(0, y_0) = |y_0|/(v_P - v_F)$ . Or, n'importe autre  $v$  tel que son cosinus ne s'annule pas avant  $t_c^*$  conduit à une capture ultérieure. En effet,

$$y(t) = y_0 + v_F \int_0^t \sin(v(\tau)) d\tau - v_P \int_0^t \sin(u(\tau)) d\tau$$

ne peut s'annuler avant  $\frac{|y_0|}{v_P - v_F}$ .

Contre la pire des perturbations, l'intervalle

$$X_{t_c} = [m(t_c) - l(t_c), m(t_c) + l(t_c)]$$

n'est donc jamais réduit à un singleton, mais puisqu'il y a capture à  $t_c$ , 0 appartient à  $X_{t_c}$ . Si  $y(\cdot)$  ne s'annule jamais avant  $t_c$ , alors il existe une autre perturbation  $\tilde{\omega}^{t_c}$  compatible avec les mêmes observations  $y^{t_c}$  telle que  $x(t_c) \neq 0$ , et qui conduit par conséquent à une capture ultérieure.  $\square$

**Proposition 4** Lorsque  $y_0 = 0$ , la valeur optimale du temps de capture est :

$$\bar{t}_c(m_0, l_0, 0) = \begin{cases} \frac{|m_0| + l_0}{v_P - v_F} & \text{si } |m_0| \geq l_0 \\ \frac{2l_0}{v_P - v_F} + \frac{v_P + v_F}{(v_P - v_F)^2} (l_0 - |m_0|) & \text{sinon} \end{cases}$$

et un contrôleur optimal  $\tilde{u}$  défini en fonction de  $t_1 = \inf\{t \mid |m(t)| = l_0\}$  par

$$\begin{cases} \sin(\tilde{u}(t)) & \stackrel{p.p.}{=} \lim_{\substack{\tau \rightarrow t \\ \tau < t}} \left\{ \sin(\tilde{u})(\tau) + \frac{\dot{y}(\tau)}{v_P} \right\} \\ \operatorname{sgn}(\cos(\tilde{u}(t))) & = \begin{cases} -\operatorname{sgn}(m_0) & |m_0| \geq l_0 \\ \left\{ \begin{array}{l} \operatorname{sgn}(m_0) \quad , \quad t < t_1 \\ -\operatorname{sgn}(m_0) \quad , \quad t \geq t_1 \end{array} \right\} & |m_0| < l_0 \end{cases} \end{cases}$$

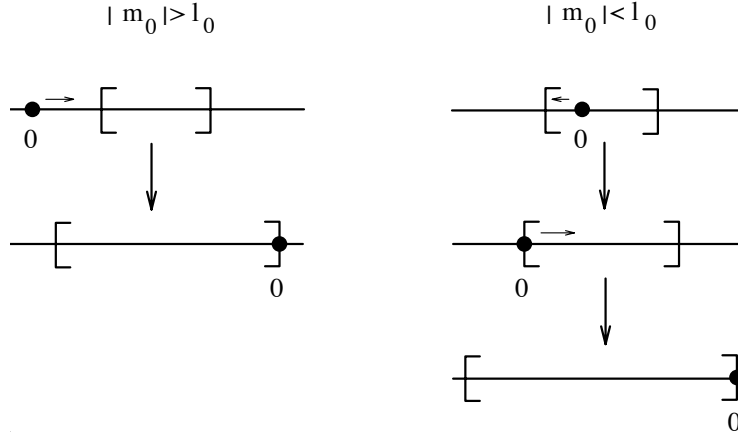


FIGURE 1 – Comportements de l'ensemble  $[m(t) - l(t), m(t) + l(t)]$  au cours du temps

### Preuve

Nous allons tout d'abord montrer que les valeurs annoncées pour  $\bar{t}_c$  sont des bornes inférieures, puisqu'elles sont atteintes pour le contrôleur proposé.

a)  $|m_0| \geq l_0$  :

Il est clair qu'un contrôleur  $\mu$  ne peut faire mieux qu'en information parfaite :

$$\bar{t}_c(m_0, l_0, 0) \geq \max_{x_0 \in X_0} t_c^*(x_0, 0) = \frac{|m_0| + l_0}{v_P - v_F}$$

b)  $|m_0| < l_0$  :

Supposons qu'il existe un contrôleur  $\mu \in \mathcal{M}$  garantissant une capture  $t_c$  en temps fini.

Soient les deux perturbations :

$$\omega^1 : \begin{cases} x_0 & = m_0 + l_0 \\ \sin(v(\cdot)) & = 0 \\ \cos(v(\cdot)) & = 1 \end{cases} \quad \omega^2 : \begin{cases} x_0 & = m_0 - l_0 \\ \sin(v(\cdot)) & = 0 \\ \cos(v(\cdot)) & = -1 \end{cases}$$

Celles-ci génèrent les mêmes observations  $y(\cdot)$ , et donc la même synthèse en boucle ouverte  $u(\cdot)$  du contrôleur  $\mu$ . Soient  $t_c^1$  et  $t_c^2$  les temps de capture respectifs. Comme  $x(t_c^1) = 0$  et  $x(t_c^2) = 0$ , on a :

$$m_0 + l_0 = v_P \int_0^{t_c^1} \cos(u(\tau)) d\tau - v_F t_c^1 \quad (2)$$

$$m_0 - l_0 = v_P \int_0^{t_c^2} \cos(u(\tau)) d\tau + v_F t_c^2 \quad (3)$$

d'où l'on déduit des bornes inférieures sur les temps de capture :

$$t_c^1 \geq \frac{|m_0| + l_0}{v_P - v_F}, \quad t_c^2 \geq \frac{l_0 - |m_0|}{v_P - v_F}$$

En retranchant (2) à (1), on obtient :

$$2l_0 = v_P \int_{t_c^2}^{t_c^1} \cos(u(\tau)) d\tau - v_F (t_c^1 + t_c^2)$$

soit :

$$2l_0 \leq v_P |t_c^2 - t_c^1| - v_F (t_c^1 + t_c^2)$$

Ainsi,

$$\max(t_c^1, t_c^2) \geq \frac{2l_0}{v_P - v_F} + \min(t_c^1, t_c^2) \frac{v_P + v_F}{v_P - v_F}$$

Si  $t_c^1 \geq t_c^2$ , on obtient :

$$t_c^1 \geq \frac{2l_0}{v_P - v_F} + \frac{v_P + v_F}{(v_P - v_F)^2} (l_0 - |m_0|)$$

Sinon,

$$t_c^2 \geq \frac{2l_0}{v_P - v_F} + \frac{v_P + v_F}{(v_P - v_F)^2} (l_0 + |m_0|)$$

Le meilleur des contrôleur  $\mu$  ne pouvant faire mieux que  $\min(t_c^1, t_c^2)$ , on déduit la borne :

$$\min_{\mu \in \mathcal{M}} \max_{\substack{(x_0, v) \\ \in \Omega(l_0, m_0)}} t_c(x_0, y_0, \mu(\cdot), v(\cdot)) \geq \frac{2l_0}{v_P - v_F} + \frac{v_P + v_F}{(v_P - v_F)^2} (l_0 - |m_0|)$$

Nous allons maintenant montrer que ces bornes sont atteintes.

Un contrôleur  $\tilde{u}(\cdot)$  comme défini dans l'énoncé garantit  $\dot{y} = 0$  presque partout et par conséquent  $y = 0$  (voir une discussion en annexe sur ce point).

**Lemme** Si  $\beta > 1$ , la fonction :

$$f(v) = \sqrt{\beta^2 - \sin(v)^2} - |\cos(v)|, \quad v \in [0, 2\pi]$$

vérifie :

$$\min_{v \in [0, 2\pi]} f(v) = f(0) = f(\pi) = \beta - 1$$

Nous laissons au lecteur le soin de vérifier cette assertion.

–  $|m_0| \geq l_0$  :

En jouant  $\tilde{u}$ , on a pour presque tout  $t$  :

$$\begin{cases} \dot{x}(t) &= v_F \cos(v(t)) - \operatorname{sgn}(m_0) \sqrt{v_P^2 - v_F^2 \sin^2(v(t))} \\ \dot{y}(t) &= 0 \end{cases}$$

D'après le lemme, quelque soit le contrôle  $v$ ,  $\operatorname{sgn}(\dot{x}(t)) = -\operatorname{sgn}(m_0)$  et la capture a donc lieu en temps fini, quelque soit  $x_0 \in X_0$ . Le pire des  $v$  minimisant  $\dot{x}$  est obtenu pour  $v(t) = 0$ ,  $\forall t$ .

$$\max_{v \in \mathcal{V}} t_c(x_0, 0, \tilde{u}(\cdot), v(\cdot)) = \frac{|x_0|}{v_P - v_F}$$

$x_0 = m_0 + \operatorname{sgn}(m_0)l_0$  est alors le pire des états initiaux pour  $u$ .

–  $|m_0| < l_0$  :

En optant pour  $\tilde{u}(\cdot)$ , il vient :

$$\begin{cases} \dot{l}(t) &= v_F |\cos(v(t))| \\ \dot{y}(t) &= 0 \end{cases}$$

La poursuite s'effectue en deux phases, mais comme les équations adjointes du problème de contrôle en  $v$  associé sont :

$$\frac{\partial \dot{t}_c}{\partial m} = 0, \quad \frac{\partial \dot{t}_c}{\partial l} = 0, \quad \frac{\partial \dot{t}_c}{\partial y} = 0$$

$v^*$  est constant durant chaque phase :  $v_1^*$  puis  $v_2^*$

– 1<sup>ère</sup> phase :  $t \leq t_1$

$$\dot{m}(t) = \operatorname{sgn}(m_0) \sqrt{v_P^2 - v_F^2 \sin^2(v(t))}$$

D'après le lemme, quelque soit le contrôle  $v(\cdot)$ ,  $|m(t_1)| = l(t_1)$  est atteint pour  $t_1 < +\infty$ .

$$t_1 = \frac{l_0 - |m_0|}{\sqrt{v_P^2 - v_F^2 \sin^2(v_1^*)} - v_F |\cos(v_1^*)|}$$

– 2<sup>ème</sup> phase :  $t > t_1$

$$\dot{m}(t) = -\operatorname{sgn}(m_1) \sqrt{v_P^2 - v_F^2 \sin^2(v(t))}$$

De même,  $\exists t_2 < +\infty$  tel que  $|m(t_2)| + l(t_2) = 0$  pour tout  $v(\cdot)$ . Alors

$$\bar{t}_c(x_0, 0, \tilde{u}(\cdot), v(\cdot)) \leq t_2$$

En effet,  $y$  est constamment nul, et :

$$\begin{aligned}\forall x \in X_{t_1}, \operatorname{sgn}(x) &= +\operatorname{sgn}(m_0) \text{ ou } x = 0 \\ \forall x \in X_{t_2}, \operatorname{sgn}(x) &= -\operatorname{sgn}(m_0) \text{ ou } x = 0\end{aligned}$$

Or,  $t \mapsto x(t)$  est continue, comme solution de  $S(x_0, 0, \tilde{u}(\cdot), v(\cdot))$ . Donc  $\exists t \leq t_2$ ,  $x(t) = 0$ .  $t_2$  vérifie :

$$t_2 = t_1 + \frac{2l(t_1)}{\sqrt{v_P^2 - v_F^2 \sin(v_2^*)^2} - v_F |\cos(v_2^*)|}$$

A  $v_1^*$  fixé, d'après le lemme,  $t_2$  est maximisé par  $v_2^* = 0$  :

$$\begin{aligned}t_2 &\leq t_1 + \frac{2l(t_1)}{v_P - v_F} \\ &\leq \frac{2l_0}{v_P - v_F} + \left( \frac{2v_F |\cos(v_1^*)|}{v_P - v_F} + 1 \right) t_1 \\ &\leq \frac{2l_0}{v_P - v_F} + \left( \frac{2v_F |\cos(v_1^*)|}{v_P - v_F} + 1 \right) \frac{l_0 - |m_0|}{\sqrt{v_P^2 - v_F^2 \sin(v_1^*)^2} - v_F |\cos(v_1^*)|}\end{aligned}$$

Nous laissons au lecteur le soin de vérifier que la fonction :

$$g(v) = \frac{\beta - 1 + 2|\cos(v)|}{(\beta - 1)f(v)}$$

possède le même tableau de variation que  $-f$ , et que par conséquent son maximum est atteint pour  $v_1^* = 0$ , d'où :

$$t_2 \leq \frac{2l_0}{v_P - v_F} + \frac{v_P + v_F}{(v_P - v_F)^2} (l_0 - |m_0|)$$

En conclusion, ce contrôleur permet de ramener  $x$  en 0 au plus tard à  $\bar{t}_c(m_0, l_0, 0)$ , en conservant la coordonnée  $y$  constamment nulle.  $\square$

Lorsque  $y_0 \neq 0$ , l'étude se ramène donc à celle d'un jeu différentiel en  $(m, l, y)$  où la cible est la variété  $\{y = 0\}$  et le critère :

$$J((m_0, l_0, y_0), u(\cdot), v(\cdot)) = K(m(t_y), l(t_y)) + \int_0^{t_y} d\tau$$

avec :

$$t_y = \inf_{t \geq 0} \{t \mid y(t) = 0\}$$

$$K(m_y, l_y) = \begin{cases} \frac{|m_y| + l_y}{v_P - v_F} & \text{si } |m_y| \geq l_y \\ \frac{2l_y}{v_P - v_F} + \frac{v_P + v_F}{(v_P - v_F)^2} (l_y - |m_y|) & \text{sinon} \end{cases}$$



et l'on a :

$$\bar{t}_c(m_0, l_0, y_0) = \min_{\phi \in \Phi} \max_{v \in \mathcal{V}} J((m_0, l_0, y_0), \phi(\cdot), v(\cdot)).$$

On se retrouve alors avec un problème classique de jeux différentiels où l'état, parfaitement connu des joueurs, est  $(m, l, y)$  au lieu de la paire  $(x, y)$ .

L'équation d'Isaacs s'écrit :

$$1 + v_F \sqrt{\left(\frac{\partial \bar{t}_c}{\partial l}\right)^2 + \left(\frac{\partial \bar{t}_c}{\partial y}\right)^2} - v_P \sqrt{\left(\frac{\partial \bar{t}_c}{\partial m}\right)^2 + \left(\frac{\partial \bar{t}_c}{\partial y}\right)^2} = 0$$

et les retours d'état optimaux sont :

$$\begin{cases} \cos(u^*) &= \frac{\partial \bar{t}_c}{\partial m} / \sqrt{\left(\frac{\partial \bar{t}_c}{\partial m}\right)^2 + \left(\frac{\partial \bar{t}_c}{\partial y}\right)^2} \\ \sin(u^*) &= \frac{\partial \bar{t}_c}{\partial y} / \sqrt{\left(\frac{\partial \bar{t}_c}{\partial m}\right)^2 + \left(\frac{\partial \bar{t}_c}{\partial y}\right)^2} \\ \cos(v^*) &= \frac{\partial \bar{t}_c}{\partial l} / \sqrt{\left(\frac{\partial \bar{t}_c}{\partial l}\right)^2 + \left(\frac{\partial \bar{t}_c}{\partial y}\right)^2} \\ \sin(v^*) &= \frac{\partial \bar{t}_c}{\partial y} / \sqrt{\left(\frac{\partial \bar{t}_c}{\partial l}\right)^2 + \left(\frac{\partial \bar{t}_c}{\partial y}\right)^2} \end{cases}$$

**Proposition 5**  $\forall (m_0, l_0, y_0) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R}^*$ , il existe une stratégie optimale pour  $u$  qui consiste en une rétroaction régulière  $\tilde{u}(m, l, y)$  jusqu'à atteindre la variété  $\{y = 0\}$  puis à choisir la stratégie explicitée dans la proposition 3.

**Preuve**

Les variables adjointes du jeu en  $(m, l, y)$  étant constantes, les extrémales sont des droites dont les pentes sont données par les conditions de transversalité. En choisissant comme paramètre du problème :

$$\delta = \frac{2v_P}{v_P - v_F} \in [2, +\infty[$$

celles-ci s'écrivent :

$$\frac{\partial \bar{t}_c}{\partial m} = \begin{cases} -\frac{(\delta - 1)\text{sgn}(m_y)}{v_P - v_F} & \text{si } |m_y| < l_y \\ \frac{\text{sgn}(m_y)}{v_P - v_F} & \text{si } |m_y| > l_y \end{cases}, \quad \frac{\partial \bar{t}_c}{\partial l} = \begin{cases} \frac{\delta + 1}{v_P - v_F} & \text{si } |m_y| < l_y \\ \frac{1}{v_P - v_F} & \text{si } |m_y| > l_y \end{cases}$$

Lorsque  $|m_y| = l_y$ , une construction de Tonelli [3] permet de rejoindre les deux champs d'extrémales, en faisant varier l'adjoint en  $(m, l)$  dans l'enveloppe convexe

des adjoints des deux zones  $\{|m_y| > l_y\}$  et  $\{|m_y| < l_y\}$ . Soit  $\lambda \in [0, 1]$  le coefficient de convexité ( $\lambda = 0$  correspond à  $|m_y| \geq l_y$ , et  $\lambda = 1$  à  $|m_y| \leq l_y$ )

$$\frac{\partial \bar{t}_c}{\partial m}(\lambda) = \frac{1 - \delta\lambda}{v_P - v_F} \text{sgn}(m_y), \quad \frac{\partial \bar{t}_c}{\partial l}(\lambda) = \frac{1 + \delta\lambda}{v_P - v_F}.$$

L'équation d'Isaacs devient :

$$2 + (\delta - 2) \sqrt{(\delta\lambda + 1)^2 + (v_P - v_F)^2} \left( \frac{\partial \bar{t}_c}{\partial y} \right)^2 = \delta \sqrt{(\delta\lambda - 1)^2 + (v_P - v_F)^2} \left( \frac{\partial \bar{t}_c}{\partial y} \right)^2$$

Or la fonction :

$$f_\lambda(z) = 2 + (\delta - 2) \sqrt{(\delta\lambda + 1)^2 + z^2} - \delta \sqrt{(\delta\lambda - 1)^2 + z^2}$$

n'admet qu'une racine positive  $r(\lambda)$ , quels que soient  $\delta > 2$  et  $\lambda \in [0, 1]$ . En effet,

$$\begin{cases} f'_\lambda(z) &= z \frac{(\delta - 2) \sqrt{(\delta\lambda + 1)^2 + z^2} - \delta \sqrt{(\delta\lambda - 1)^2 + z^2}}{2 \sqrt{((\delta\lambda + 1)^2 + z^2)((\delta\lambda - 1)^2 + z^2)}} \leq 0 \\ f_\lambda(0) &= 2\delta(1 - \lambda) \geq 0 \text{ si } (\delta\lambda - 1) \geq 0, \quad 2\delta\lambda(\delta - 1) \geq 0 \text{ sinon} \\ \lim_{z \rightarrow +\infty} f_\lambda(z) &= -\infty \end{cases}$$

Ainsi, l'adjoint en  $y$  est déterminé de manière unique par l'expression :

$$\frac{\partial \bar{t}_c}{\partial y}(\lambda) = \text{sgn}(y_0) \frac{r(\lambda)}{v_P - v_F}$$

On constate que  $r(0) = 0$  et  $r(1) = 0$  sont racines de  $f$ , et donc que les extrémales associées restent dans le plan  $y = 0$ . Les autres extrémales ( $\lambda \in ]0, 1[$ ) arrivent donc nécessairement sur  $y = 0$  en  $|m_y| = l_y$ , à l'instant  $t_y$  :

$$t_y(\lambda, y_0) = \frac{|y_0|}{\left( \frac{v_P}{\sqrt{(1 - \delta\lambda)^2 + r(\lambda)^2}} - \frac{v_F}{\sqrt{(1 + \delta\lambda)^2 + r(\lambda)^2}} \right) r(\lambda)}$$

avec :

$$\begin{cases} m_y(\lambda) &= m_0 - v_P \frac{(1 - \delta\lambda) \text{sgn}(m_y)}{\sqrt{(1 - \delta\lambda)^2 + r(\lambda)^2}} t_y(\lambda, y_0) \\ l_y(\lambda) &= l_0 + v_F \frac{(1 + \delta\lambda)}{\sqrt{(1 + \delta\lambda)^2 + r(\lambda)^2}} t_y(\lambda, y_0) \end{cases}$$

Le choix du signe de  $m_y$  est à l'initiative du minimiseur  $u$ , et pour une condition initiale  $(m_0, l_0, y_0)$ ,  $\lambda$  doit vérifier  $|m_y| = l_y$  soit  $\text{sgn}(m_y) m_0 = l_0 + e(\lambda) |y_0|$  où  $e$  est la fonction :

$$e(\lambda) = \frac{\delta \sqrt{(1 + \delta\lambda)^2 + r(\lambda)^2} + (\delta - 2) \sqrt{(1 - \delta\lambda)^2 + r(\lambda)^2}}{\left( \delta \sqrt{(1 + \delta\lambda)^2 + r(\lambda)^2} - (\delta - 2) \sqrt{(1 - \delta\lambda)^2 + r(\lambda)^2} \right) r(\lambda)} - \frac{\delta\lambda}{r(\lambda)}$$

En  $y_0 \neq 0$ ,  $\lambda$  est différent de 0 et de 1, et par conséquent  $r(\lambda) \neq 0$ . D'autre part, sur  $]0, 1[ \times ]0, +\infty[$ , la fonction  $(\lambda, z) \mapsto f_\lambda(z)$  est  $C^\infty$ , et  $f'_\lambda(z) \neq 0$ . En accord avec le théorème des fonctions implicites,  $r(\cdot)$  est donc  $C^1$  sur  $]0, 1[$ . Comme  $r(\cdot)$  ne s'annule pas, la fonction  $\lambda \mapsto e(\lambda)$  est également  $C^1$  par composition sur  $]0, 1[$ . Or,

$$\begin{cases} \lim_{\lambda \rightarrow 0} e(\lambda) = \lim_{\lambda \rightarrow 0} \frac{\delta(1-\lambda) - 1}{r(\lambda)} = +\infty \\ \lim_{\lambda \rightarrow 1} e(\lambda) = \lim_{\lambda \rightarrow 1} \frac{1 - \delta^2}{(2\delta - 1)r(\lambda)} = -\infty \end{cases}$$

Ainsi  $\text{Im}(e) = \mathbb{R}$ , et par conséquent, quelque soit le triplet  $(m, l, y \neq 0)$ , il existe au moins un  $\tilde{\lambda}(m, l, y) \in ]0, 1[$  solution de :

$$e(\tilde{\lambda}) = \frac{\text{sgn}(m_y)m - l}{|y|}$$

tel que le contrôle  $\tilde{u}(m, l, y)$  :

$$\cos(\tilde{u}) = \frac{(1 - \delta\tilde{\lambda})\text{sgn}(m_y)}{\sqrt{(1 - \delta\tilde{\lambda})^2 + r(\tilde{\lambda})^2}}, \quad \sin(\tilde{u}) = \frac{r(\tilde{\lambda})\text{sgn}(y)}{\sqrt{(1 - \delta\tilde{\lambda})^2 + r(\tilde{\lambda})^2}}$$

soit une rétroaction optimale pour rejoindre  $(m_y, |m_y|, 0)$  au plus tard au temps  $t_y(\tilde{\lambda}(m, l, y), y)$ .  
□

Des calculs plus poussés permettent, à l'aide du calcul formel, d'obtenir une expression analytique de  $r$  :

$$r(\lambda) = \left( \frac{\delta(\delta - 2) [\sqrt{1 - 4\lambda\delta + 4\lambda\delta^2} + 1]}{2(\delta - 1)^2} + \frac{\lambda\delta^3}{\delta - 1} + \frac{1}{(\delta - 1)^2} - (\lambda\delta + 1)^2 \right)^{1/2}$$

et de tracer numériquement le graphe de la fonction  $e$  :

**Conjecture**  $\forall \delta \in ]2, +\infty[$ , la fonction  $e$  est strictement décroissante sur  $]0, 1[$ .

**Proposition 6** *Sous réserve que la conjecture soit vérifiée,*

- $\forall (m, l, y) \in \mathbb{R}^* \times \mathbb{R}^+ \times \mathbb{R}^*$ ,  $\tilde{u}$  est unique et  $\text{sgn}(m_y) = \text{sgn}(m)$ .
- $\{0\} \times \mathbb{R}^+ \times \mathbb{R}^*$  est un plan de dispersion où le signe de  $m_y$  est à l'initiative du minimiseur  $u$ , conduisant à des trajectoires optimales symétriques par rapport au plan  $\{m = 0\}$  dans l'espace  $(m, l, y)$ .

**Preuve**

Si  $e'(\cdot)$  ne s'annule jamais, alors pour tout triplet  $(m, l, y \neq 0)$ , il existe un unique  $\tilde{\lambda}$  solution de  $e(\lambda) = (|m| - l)/|y|$ , et de plus la fonction  $(m, l, y) \mapsto \tilde{\lambda}(m, l, y)$  est  $C^1$  sur  $\mathbb{R}^* \times \mathbb{R}^+ \times \mathbb{R}^*$ , à nouveau par le théorème des fonctions implicites. Contre le pire des  $v(\cdot)$ ,  $\tilde{\lambda}$  est constant, et pour  $\text{sgn}(m_y) = \text{sgn}(m)$ , le temps final correspondant est :

$$\bar{t}_c(m, l, y) = \frac{2|m_y(\tilde{\lambda}(m, l, y), m, y)|}{v_P - v_F} + t_y(\tilde{\lambda}(m, l, y), y)$$

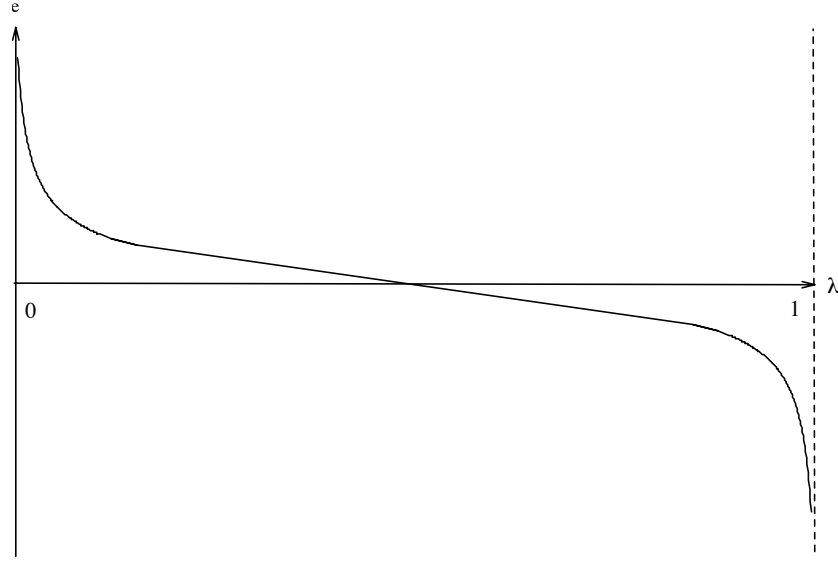


FIGURE 2 – La fonction  $e(\lambda)$

Par construction,  $\bar{t}_c$  est une solution  $C^1$  de l'équation d'Isaacs sur  $\mathbb{R}^* \times \mathbb{R}^+ \times \mathbb{R}^*$ , dont le gradient est :

$$\begin{cases} \frac{\partial \bar{t}_c}{\partial m}(m, l, y) = \frac{1 - \delta \tilde{\lambda}(m, l, y)}{v_P - v_F} \operatorname{sgn}(m) \\ \frac{\partial \bar{t}_c}{\partial l}(m, l, y) = \frac{1 + \delta \tilde{\lambda}(m, l, y)}{v_P - v_F} \\ \frac{\partial \bar{t}_c}{\partial y}(m, l, y) = \frac{r(\tilde{\lambda}(m, l, y))}{v_P - v_F} \operatorname{sgn}(y) \end{cases}$$

Il reste à vérifier qu'elle est également solution des équations de viscosité [4] aux points de non-différentiabilité, c'est à dire en  $m = 0$ .

Or,  $\forall (l, y) \in \mathbb{R}^+ \times \mathbb{R}^*$ ,  $1 - \delta \tilde{\lambda}(0, l, y) < 0$ . En effet,  
 $e(\cdot)$  est décroissant  
 $e(\lambda) = 0 \Rightarrow 1 - \delta \lambda < 0$   
 $e(\tilde{\lambda}(0, l, y)) = -\frac{l}{|y|} < 0$

En posant  $\frac{1 - \delta\tilde{\lambda}(0, l, y)}{v_P - v_F} = -\alpha < 0$ , les sous et sur-gradient de  $\bar{t}_c$  s'écrivent :

$$\begin{cases} \partial^- \bar{t}_c(0, l, y) &= \emptyset \times \left\{ \frac{\partial \bar{t}_c}{\partial l}(0, l, y) \right\} \times \left\{ \frac{\partial \bar{t}_c}{\partial y}(0, l, y) \right\} \\ \partial^+ \bar{t}_c(0, l, y) &= [-\alpha, \alpha] \times \left\{ \frac{\partial \bar{t}_c}{\partial l}(0, l, y) \right\} \times \left\{ \frac{\partial \bar{t}_c}{\partial y}(0, l, y) \right\} \end{cases}$$

Il suffit donc de vérifier que le candidat  $\bar{t}_c(\cdot)$  est une sur-solution de viscosité de l'équation d'Isaacs, c'est-à-dire :

$$\forall \rho \in \partial^+ \bar{t}_c(0, l, y), \quad -1 + H^*((0, l, y), -\rho) \leq 0$$

ce qui est trivialement le cas puisque :

$$\begin{aligned} 1 + v_F \sqrt{\left(\frac{\partial \bar{t}_c}{\partial l}\right)^2 + \left(\frac{\partial \bar{t}_c}{\partial y}\right)^2} - v_P \sqrt{\rho_m^2 + \left(\frac{\partial \bar{t}_c}{\partial y}\right)^2} &\geq \\ 1 + v_F \sqrt{\left(\frac{\partial \bar{t}_c}{\partial l}\right)^2 + \left(\frac{\partial \bar{t}_c}{\partial y}\right)^2} - v_P \sqrt{\alpha^2 + \left(\frac{\partial \bar{t}_c}{\partial y}\right)^2} &= 0 \end{aligned}$$

□

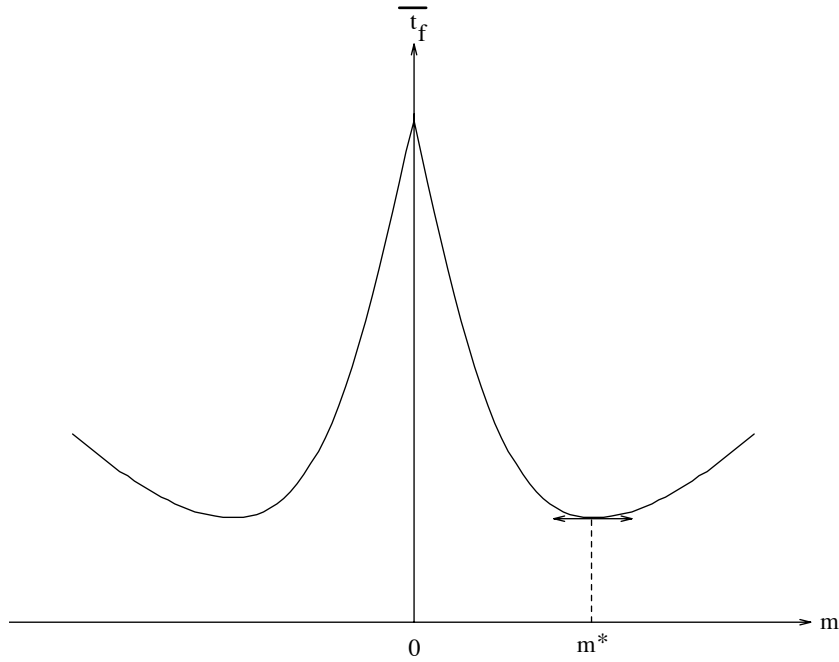


FIGURE 3 –  $\bar{t}_c$  fonction de  $m$  à  $(l, y)$  fixés

## 4 Conclusion

On a pu étendre au jeu, ou plutôt à la commande minimax, en information imparfaite à temps final variable la théorie développée précédemment pour le temps final fixé. Cette extension a été facilitée par le fait que la capture est directement observable par le joueur qui manque d'information, ici le poursuivant. Dans le cas contraire, il faut recourir à une inéquation variationnelle à la place de l'équation d'Isaacs.

Cette théorie apporte une solution complète au jeu considéré. C'est, à notre connaissance, la première application non banale de cette théorie publiée. Si la solution trouvée est intuitivement raisonnable, nous n'en connaissons pas de preuve d'optimalité en dehors de cette théorie.

On peut, à l'aide de cette même théorie, résoudre diverses variantes du même jeu : capture donnée par une distance euclidienne inférieure ou égale à  $\ell$  donné, ou par un "ensemble de capture" d'une autre forme, ou encore distance minimax à un instant final donné.

## 5 Annexe

Les contrôles  $\mu$  et  $v$  envisagés fournissent existence et unicité d'une trajectoire solution du système  $S$ , mais pas le caractère régulier de  $\dot{y}$ . En effet,  $v(\cdot)$  n'étant supposée être qu'une fonction mesurable du temps,  $\dot{y}$  peut a priori être discontinue. Néanmoins, par le théorème de Lusin, n'importe quelle fonction mesurable peut être rendue continue, en ôtant de son domaine de définition un ensemble de points de mesure arbitrairement petite. Comme  $u$  n'a accès à l'instant  $t$  qu'à des mesures de  $y(\tau)$  pour  $\tau \leq t$ , il ne peut en déduire qu'une limite à gauche de  $\dot{y}$  en  $t$ , et ne mesurer réellement  $\dot{y}(t)$  que là où il est continu.

Lorsque  $v(\cdot)$  est continue par morceaux, le contrôleur  $\tilde{u}$  de la proposition 3 :

$$\sin(\tilde{u}(t)) = \lim_{\substack{p.p. \\ \tau \rightarrow t \\ \tau < t}} \left\{ \sin(\tilde{u}(\tau)) + \frac{\dot{y}(\tau)}{v_P} \right\}$$

est bien défini et garantit  $\dot{y} = 0$  presque partout.

Si  $v(\cdot)$  n'est pas continue par morceaux,  $\tilde{u}$  n'est pas défini partout, mais garantit  $\dot{y} = 0$  excepté sur un ensemble arbitrairement petit.

Comme  $|\dot{y}|$  est bornée par  $v_P + v_F$ , alors pour tout  $v(\cdot) \in \mathcal{V}$  et pour tout  $\epsilon > 0$ , la synthèse en boucle ouverte du contrôleur  $\tilde{u}$  peut donc être prolongée de telle sorte que  $|y(t)| \leq \epsilon(v_p + v_F)t$  pour tout  $t$ . P peut ainsi se ramener arbitrairement proche de F.

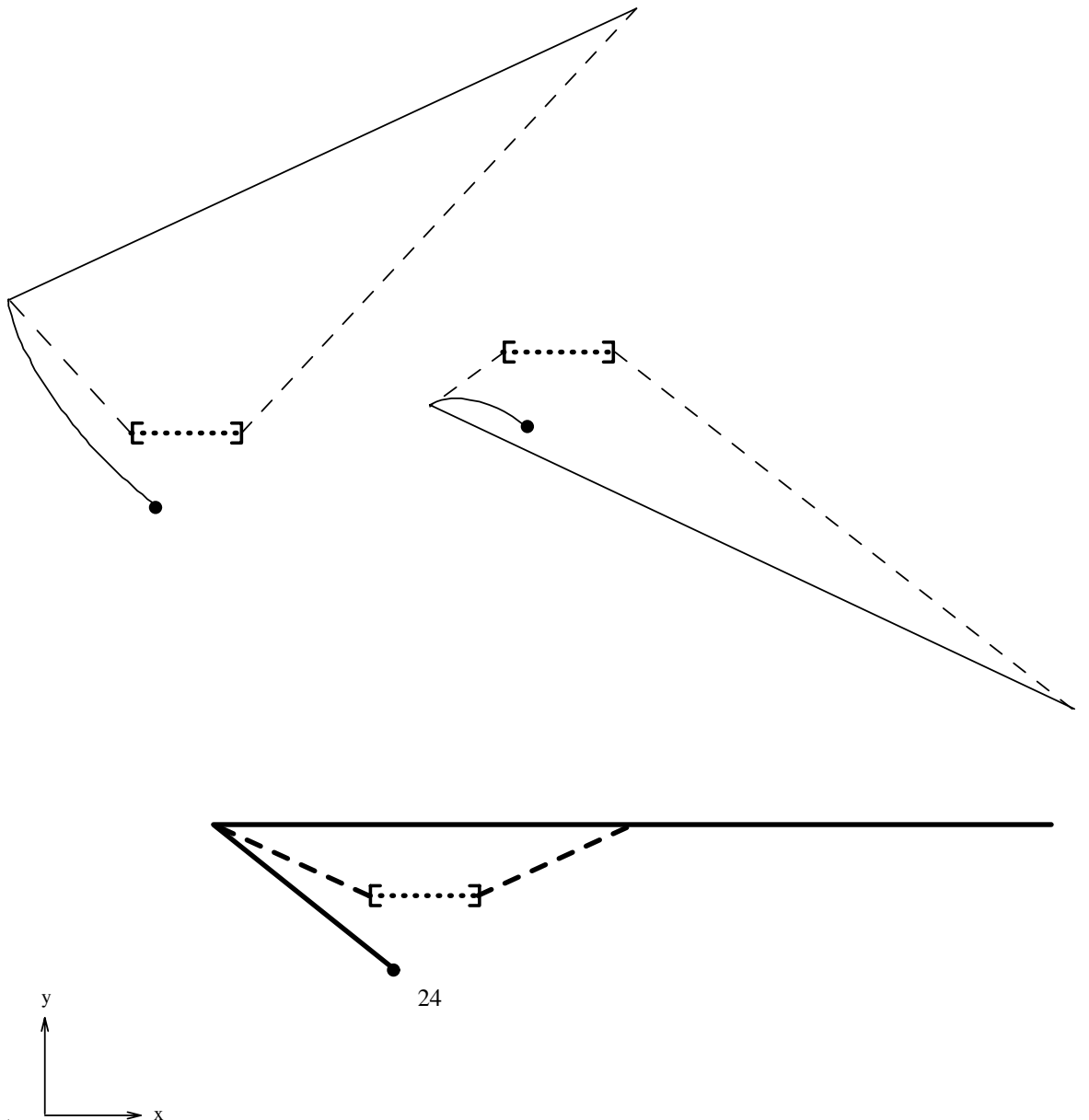
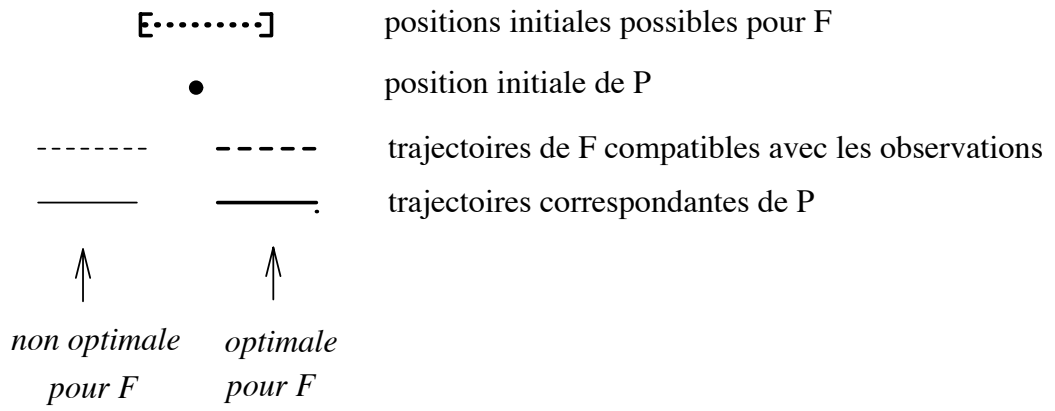


FIGURE 4 – Stratégie optimale de P contre divers comportements de F



## Références

- [1] R.ISAACS. *Differential Games*. Wiley, 1965.
- [2] P.BERNHARD. *Sketch of a Theory of Nonlinear Partial Information Min-Max Control*. Rapport de Recherche INRIA 2020, August 1993.
- [3] C.CARATHÉODORY. *Calculus of Variations and Partial Differential Equations of the First Order*, chapter 15, pages 276–277. Chelsea Publishing Company (réimpression), 1935.
- [4] W.H. FLEMING AND H.M. SONER. *Controlled Markov Processes and Viscosity Solutions*. Springer-Verlag, Applications of Mathematics 25, 1993.
- [5] M.R. James, J.S. Baras, and R.J. Elliott. Output feedback risk-sensitive control and differential games for continuous-time nonlinear systems. In *Proceedings of the 32nd Conference on Decision and Control*, pages 3357–3360, San Antonio, Texas, December 1993.