Correction détaillée du sujet d'examen CC no. 2
MAM5-INUM – Commande optimale

1 Exercice 1 (15 points)

On considère une barque se déplaçant à vitesse constante dans un canal rectiligne, de largeur constante. On peut supposer que la norme de la vitesse vaut 1, et normaliser de même la largeur du canal à 1. Il y a dans ce canal un courant $c(y)$ dirigé selon (Ox) que l'on suppose fort (avec $c \in C^\infty$), c'est-à-dire tel que $c(y) > 1$, pour tout y.
On contrôle directement le cap du bateau, noté $u(t)$, de sorte que le système s'écrit

$$\begin{cases} \dot{x}(t) = \cos u(t) + c\big(y(t)\big), \\ \dot{y}(t) = \sin u(t), \end{cases}$$

avec $(x, y) \in \mathbb{R}^2$ les coordonnées dans le canal (voir figure ci-après) et $u \in \mathbb{R}$. Dans tout l'exercice on prend une condition initiale à l'origine,

$$(x(0), y(0)) = (0, 0),$$

le temps final t_f est supposé libre, et on admettra que tous les problèmes considérés possèdent une solution.

1.1 Partie A – Minimisation du déport $x(t_f)$

On s'intéresse pour commencer à minimiser le déport $x(tf)$ quand on vise la berge opposée, $y(tf) = 1$.

1.1.1 1.1 Forme de Lagrange

Énoncé. Mettre le problème sous forme de Lagrange avec un intégrande f^0 que l'on précisera.

Correction. Le problème consiste à minimiser $x(t_f)$ avec les conditions :
- Conditions initiales : $x(0) = 0$, $y(0) = 0$;
- Conditions finales : $y(t_f) = 1$, $x(t_f)$ libre ;
- Temps final t_f libre.

La fonction coût s'écrit :

$$J(u) = x(t_f) = \int_0^{t_f} \dot{x}(t)\, dt = \int_0^{t_f} \big[\cos u(t) + c(y(t))\big]\, dt.$$

Donc l'intégrande est :

$$f^0(x, y, u) = \cos u + c(y).$$

Explication. On exprime le coût final comme l'intégrale de la dérivée, ce qui permet de se ramener à un problème de Lagrange standard.

1.1.2 1.2 Hamiltonien

Énoncé. Écrire le Hamiltonien du problème.

Correction. Le Hamiltonien général pour un problème de Lagrange est

$$H = p^0 f^0 + p_x \dot{x} + p_y \dot{y},$$

En substituant les expressions de $\dot{x}$ et $\dot{y}$, on obtient :

$$H(x, y, p_x, p_y, u) = p^0 f^0 + p_x(\cos u + c(y)) + p_y \sin u,$$

où $p^0 \leq 0$ est la variable adjointe associée au coût.

Explication. Le Hamiltonien est la somme des produits des variables adjointes avec les dérivées des états, plus le terme de coût pondéré par p^0.

1.1.3 1.3 Système adjoint

Énoncé. Écrire le système adjoint.

Correction. Les équations adjointes sont données par

$$\dot{p}_x = -\frac{\partial H}{\partial x}, \qquad \dot{p}_y = -\frac{\partial H}{\partial y}.$$

On a $\frac{\partial H}{\partial x} = 0$ car H ne dépend pas explicitement de x. Donc $\dot{p}_x = 0$.
De plus,

$$\frac{\partial H}{\partial y} = (p^0 + p_x)c'(y),$$

donc $\dot{p}_y = -(p^0 + p_x)c'(y)$.
Ainsi, le système adjoint est :

$$\dot{p}_x = 0, \qquad \dot{p}_y = -(p^0 + p_x)c'(y).$$

Explication : Les équations adjointes découlent du principe du maximum de Pontryagin. On remarque que p_x est constant, ce qui reflète l'invariance par translation en x. La variation de p_y dépend du gradient du courant $c'(y)$.

1.1.4 1.4 Conditions de transversalité

Énoncé. Écrire les conditions de transversalité.

Correction.
- Pour x : $x(t_f)$ libre donc $p_x(t_f) = 0$;
- Pour y : $y(t_f) = 1$ fixé donc pas de condition sur $p_y(t_f)$;
- Pour t_f libre : $H(t_f) = 0$.

Explication. Les conditions de transversalité lient les variables adjointes finales aux contraintes finales sur l'état.

1.1.5 1.5 Cas normal

Énoncé. Montrer qu'on est nécessairement dans le cas normal.

Correction. Supposons par l'absurde que nous sommes dans le cas anormal, c'est-à-dire que $p^0 = 0$.

D'après l'équation adjointe, nous avons

$$\dot{p}_x = -\frac{\partial H}{\partial x} = 0,$$

ce qui implique que p_x est constante. La condition de transversalité nous donne $p_x(t_f) = 0$ puisque $x(t_f)$ est libre. Par conséquent,

$$p_x(t) = 0 \quad \text{pour tout } t \in [0, t_f].$$

Avec $p^0 = 0$ et $p_x = 0$, le Hamiltonien se simplifie en

$$H = p_y \sin u.$$

Nous cherchons à maximiser $H = p_y \sin u$ par rapport à u. Plusieurs cas se présentent :

— Si $p_y > 0$, le maximum est atteint lorsque $\sin u = 1$, donnant

$$H = p_y = |p_y|.$$

— Si $p_y < 0$, le maximum est atteint lorsque $\sin u = -1$, donnant

$$H = -p_y = |p_y|.$$

— Si $p_y = 0$, alors $H = 0$ quelle que soit la valeur de u.

On a alors que le maximum de H vaut $p_y * sgn(p_y) = |p_y|$ lorsque $p_y \neq 0$, et 0 lorsque $p_y = 0$.

Le temps final t_f étant libre, le principe du maximum impose

$$H(t_f) = 0.$$

Si $p_y \neq 0$, cela entraînerait

$$|p_y(t_f)| = 0,$$

ce qui est impossible. Nous devons donc avoir

$$p_y(t_f) = 0.$$

L'équation adjointe pour p_y est

$$\dot{p}_y = -(p^0 + p_x)c'(y).$$

Avec $p^0 = 0$ et $p_x = 0$, nous obtenons

$$\dot{p}_y = 0,$$

donc p_y est constante. Puisque $p_y(t_f) = 0$, il s'ensuit que

$$p_y(t) = 0 \quad \text{pour tout } t \in [0, t_f].$$

Nous aboutissons à la situation où

$$p^0 = 0, \qquad p_x(t) = 0, \qquad p_y(t) = 0 \quad \text{pour tout } t,$$

ce qui contredit le principe du maximum de Pontryagin qui exige que les multiplicateurs ne soient pas tous nuls.

L'hypothèse du cas anormal conduit à une contradiction. Le problème est donc nécessairement dans le cas normal, et nous pouvons normaliser

$$p^0 = -1.$$

Explication. Le cas anormal mène à une contradiction via la condition $H(t_f) = 0$, donc le problème est nécessairement normal.

1.1.6 1.6 Condition de maximisation

Énoncé. Appliquer la condition de maximisation.

Correction. Avec $p^0 = -1$ et $p_x = 0$, le Hamiltonien

$$H(x, y, p_x, p_y, u) = p^0 f^0 + p_x(\cos u + c(y)) + p_y \sin u,$$

avec $f^0(x, y, u) = \cos u + c(y)$, devient :

$$H = -\cos u - c(y) + p_y \sin u.$$

On cherche à maximiser H par rapport à u. On écrit :

$$H = -c(y) + \big[-\cos u + p_y \sin u\big].$$

En utilisant la formule d'addition pour le cosinus nous avons :

$$A\cos(u - \alpha) = A\cos u \cos\alpha + A\sin u \sin\alpha.$$

En identifiant avec $-\cos u + p_y \sin u$, nous obtenons :

$$A\cos\alpha = -1, \qquad A\sin\alpha = p_y.$$

En élevant au carré et en additionnant, nous trouvons :

$$A^2(\cos^2\alpha + \sin^2\alpha) = (-1)^2 + p_y^2 \ \Rightarrow\ A^2 = 1 + p_y^2 \ \Rightarrow\ A = \sqrt{1 + p_y^2}.$$

On multiplie les deux côtés par $\sqrt{1 + p_y^2}$:

$$\sqrt{1 + p_y^2}\,\cos(u - \alpha) = \sqrt{1 + p_y^2}\,\cos u \cos\alpha + \sqrt{1 + p_y^2}\,\sin u \sin\alpha.$$

Si l'on pose

$$\cos\alpha = \frac{-1}{\sqrt{1 + p_y^2}}, \qquad \sin\alpha = \frac{p_y}{\sqrt{1 + p_y^2}},$$

alors on obtient bien

$$\sqrt{1 + p_y^2}\,\cos(u - \alpha) = -\cos u + p_y \sin u.$$

Pour trouver α, considérons le nombre complexe

$$z = -1 + ip_y.$$

Son module est

$$|z| = \sqrt{(-1)^2 + p_y^2} = \sqrt{1 + p_y^2}.$$

Ainsi, on peut écrire, si $\alpha = arg(z)$:

$$\cos\alpha = \frac{\Re(z)}{|z|} = \frac{-1}{\sqrt{1 + p_y^2}}, \qquad \sin\alpha = \frac{\Im(z)}{|z|} = \frac{p_y}{\sqrt{1 + p_y^2}}.$$

L'angle α est donc exactement l'argument du nombre complexe $z = -1 + ip_y$:

$$\alpha = \arg(-1 + ip_y).$$

Le terme entre crochets peut donc s'écrire :

$$-\cos u + p_y \sin u = \sqrt{1+p_y^2}\,\cos(u-\alpha),$$

où $\alpha = \arg(-1 + ip_y)$.

On peut réecrire H :

$$H = -c(y) + \sqrt{1+p_y^2}\,\cos(u-\alpha).$$

Le maximum est atteint pour $cos(u-\alpha) = 1$, donc si :

$$u(t) = \alpha = \arg(-1 + ip_y(t)).$$

Explication. On utilise une identité trigonométrique pour exprimer la combinaison linéaire de cos et sin comme un cosinus déphasé.

1.1.7 1.7 Relation avec temps final libre

Énoncé. En utilisant le fait que le temps final est libre, montrer que $\cos u(t) = -1/c(y(t))$.

Correction. Avec u optimal, on a :

$$\cos u = \frac{-1}{\sqrt{1+p_y^2}}, \qquad \sin u = \frac{p_y}{\sqrt{1+p_y^2}}.$$

La condition $H(t) = 0$ (temps libre) donne :

$$-\cos u - c(y) + p_y \sin u = 0,$$

d'où

$$-\left(\frac{-1}{\sqrt{1+p_y^2}}\right) - c(y) + p_y\left(\frac{p_y}{\sqrt{1+p_y^2}}\right) = 0,$$

soit

$$\frac{1+p_y^2}{\sqrt{1+p_y^2}} - c(y) = 0 \quad\Longrightarrow\quad \sqrt{1+p_y^2} = c(y).$$

Donc

$$\cos u = \frac{-1}{c(y)}.$$

Bonus : on peut également exprimer $\sin u$ directement :

$$\sin u = \frac{p_y}{\sqrt{1+p_y^2}} = \frac{\sqrt{c^2(y)-1}}{c(y)}.$$

Cependant, pour déterminer le signe de $\sin u$, on utilise l'équation $\dot{y} = \sin u$. Comme on doit atteindre $y(t_f) = 1$ à partir de $y(0) = 0$, on a nécessairement $\dot{y} > 0$, donc $\sin u > 0$. On vérifie que cette condition est cohérente avec l'expression ci-dessus :

$$\sin^2 u = 1 - \cos^2 u = 1 - \frac{1}{c^2(y)} = \frac{c^2(y)-1}{c^2(y)} > 0 \quad \text{puisque } c(y) > 1,$$

d'où

$$\sin u = \frac{\sqrt{c^2(y)-1}}{c(y)} > 0.$$

Explication. La condition $H = 0$ permet d'éliminer p_y et d'exprimer le contrôle optimal directement en fonction de l'état.

1.1.8 1.8 Temps final

Énoncé. Donner l'expression du temps final.

Correction. On a

$$\dot{y} = \sin u = \frac{\sqrt{c^2(y) - 1}}{c(y)},$$

donc

$$dt = \frac{c(y)}{\sqrt{c^2(y) - 1}}\, dy,$$

et donc

$$t_f = \int_0^{t_f} dt = \int_0^1 \frac{c(y)}{\sqrt{c^2(y) - 1}}\, dy.$$

Explication. On intègre l'équation différentielle pour y entre les conditions initiale et finale.

1.2 Partie B – Temps minimal pour atteindre l'autre berge

1.2.1 1.9 Hamiltonien

Énoncé. Écrire le Hamiltonien du problème.

Correction. Pour un problème de temps minimal, $f^0 = 1$, donc :

$$H = p^0 + p_x(\cos u + c(y)) + p_y \sin u.$$

Explication. Le Hamiltonien inclut maintenant le terme p^0 correspondant au coût temporel.

1.2.2 1.10 Cas normal

Énoncé.

Énoncé. Montrer qu'on est nécessairement dans le cas normal.

Correction. Supposons par l'absurde que nous sommes dans le cas anormal, c'est-à-dire que $p^0 = 0$.
D'après l'équation adjointe, nous avons

$$\dot{p}_x = -\frac{\partial H}{\partial x} = 0,$$

ce qui implique que p_x est constante. La condition de transversalité nous donne $p_x(t_f) = 0$ puisque $x(t_f)$ est libre. Par conséquent,

$$p_x(t) = 0 \quad \text{pour tout } t \in [0, t_f].$$

Avec $p^0 = 0$ et $p_x = 0$, le Hamiltonien se simplifie en

$$H = p_y \sin u.$$

Nous cherchons à maximiser $H = p_y \sin u$ par rapport à u. Plusieurs cas se présentent :

— Si $p_y > 0$, le maximum est atteint lorsque $\sin u = 1$, donnant $H = p_y = |p_y|$.
— Si $p_y < 0$, le maximum est atteint lorsque $\sin u = -1$, donnant $H = -p_y = |p_y|$.
— Si $p_y = 0$, alors $H = 0$ quelle que soit la valeur de u.

Ainsi, le maximum de H vaut $|p_y|$ lorsque $p_y \neq 0$, et 0 lorsque $p_y = 0$.
Le temps final t_f étant libre, le principe du maximum impose $H(t_f) = 0$. Si $p_y \neq 0$, cela entraînerait $|p_y(t_f)| = 0$, ce qui est impossible. Nous devons donc avoir $p_y(t_f) = 0$.
L'équation adjointe pour p_y est

$$\dot{p}_y = -(p^0 + p_x)c'(y).$$

Avec $p^0 = 0$ et $p_x = 0$, nous obtenons $\dot{p}_y = 0$, donc p_y est constante. Puisque $p_y(t_f) = 0$, il s'ensuit que

$$p_y(t) = 0 \quad \text{pour tout } t \in [0, t_f].$$

Nous aboutissons à la situation où $p^0 = 0$, $p_x(t) = 0$ et $p_y(t) = 0$ pour tout t, ce qui contredit le principe du maximum de Pontryagin, qui exige que les multiplicateurs ne soient pas tous nuls.
L'hypothèse du cas anormal conduit donc à une contradiction. Le problème est nécessairement dans le cas normal, et nous pouvons normaliser $p^0 = -1$.

Explication. La structure du problème est similaire à celle de la partie A, conduisant à la même conclusion sur la normalité. Le cas anormal mène à une contradiction via la condition $H(t_f) = 0$, donc le problème est nécessairement normal.

1.2.3 1.11 Non-annulation de p_y

Énoncé. Montrer que p_y ne s'annule pas.

Correction. Le système adjoint donne pour l'Hamiltonien $H = p^0 + p_x(\cos u + c(y)) + p_y \sin u$:

$$\dot{p}_y = -p_x c'(y) = 0$$

(car $p_x = 0$), donc p_y est constante.
Si $p_y = 0$, alors $H = -1$ (avec $p^0 = -1$), ce qui contredit $H(t_f) = 0$.
Donc $p_y \neq 0$.

Explication. La condition $H(t_f) = 0$ et la constance de p_y empêchent p_y d'être nul.

1.2.4 1.12 Contrôle optimal

Énoncé. Appliquer la condition de maximisation.

Correction. Avec $p^0 = -1$, $p_x = 0$, le Hamiltonien est :

$$H = -1 + p_y \sin u.$$

La maximisation par rapport à u donne :
— Si $p_y > 0$, le maximum est atteint pour $\sin u = 1$, soit $u = \pi/2$;
— Si $p_y < 0$, le maximum est atteint pour $\sin u = -1$, soit $u = -\pi/2$.

Comme on veut $\dot{y} = \sin u > 0$ pour atteindre $y = 1$, on choisit $u = \pi/2$.

Explication. La maximisation est simple car le Hamiltonien dépend linéairement de $\sin u$.

1.2.5 1.13 Temps minimal

Énoncé. Déterminer le temps minimal.

Correction. Avec $u = \pi/2$, on a $\dot{y} = \sin(\pi/2) = 1$, donc

$$y(t) = t + C.$$

La condition initiale $y(0) = 0$ implique $C = 0$ et donc $y(t) = t$.

Donc pour atteindre $y = 1$, il faut donc $t_f = 1$.

Explication. La trajectoire est verticale directe, le temps minimal est donc la distance verticale divisée par la vitesse verticale.

1.3 Partie C – Temps minimal avec point d'arrivée fixé

1.3.1 1.14 Cas normal

Énoncé. Dans le cas normal, donner l'expression de u en fonction des adjoints.

Correction. Avec $p^0 = -1$, le Hamiltonien est :

$$H = -1 + p_x(\cos u + c(y)) + p_y \sin u.$$

La condition de maximisation donne

$$\frac{\partial H}{\partial u} = -p_x \sin u + p_y \cos u = 0.$$

Si $p_x \neq 0$, on en déduit

$$p_y \cos u = p_x \sin u \quad \Rightarrow \quad \tan u = \frac{p_y}{p_x}.$$

Les solutions sont donc

$$u = \arctan\left(\frac{p_y}{p_x}\right) + k\pi, \quad k \in \mathbb{Z}.$$

Par ailleurs, le terme à maximiser dans le Hamiltonien est

$$p_x \cos u + p_y \sin u,$$

qui est maximal lorsque le vecteur $(\cos u, \sin u)$ est colinéaire et de même sens que (p_x, p_y). On obtient alors

$$u = \arg(p) = \arg(p_x, p_y) = \arg(p_x + ip_y) \pmod{2\pi}.$$

La condition $H = 0$ donne :

$$-1 + p_x(\cos u + c(y)) + p_y \sin u = 0.$$

Or pour u $= \arg(p_x + ip_y)$ on a

$$\cos u = \frac{p_x}{\sqrt{p_x^2 + p_y^2}}, \qquad \sin u = \frac{p_y}{\sqrt{p_x^2 + p_y^2}}.$$

Donc

$$p_x \cos u + p_y \sin u = \sqrt{p_x^2 + p_y^2},$$

d'où

$$-1 + \sqrt{p_x^2 + p_y^2} + p_x c(y) = 0.$$

Donc $\sqrt{p_x^2 + p_y^2} = 1 - p_x c(y)$ On en déduit alors, en utilisant

$$\cos u = \frac{p_x}{\sqrt{p_x^2 + p_y^2}}, \quad \sin u = \frac{p_y}{\sqrt{p_x^2 + p_y^2}}$$

que

$$\cos u = \frac{p_x}{1 - p_x c(y)}, \quad \sin u = \frac{p_y}{1 - p_x c(y)}.$$

Explication. La condition de maximisation et $H = 0$ permettent d'exprimer le contrôle en fonction des adjoints.

1.3.2 1.15 Cas anormal

Énoncé. Dans le cas anormal, donner l'expression de u.

Correction. Dans le cas anormal, on a $p^0 = 0$. Le Hamiltonien devient :

$$H = p_x\big(\cos u + c(y)\big) + p_y \sin u.$$

La condition de maximisation par rapport à u est identique au cas normal et donne :

$$u = \arg\big(p_x + ip_y\big).$$

La condition $H = 0$ impose :

$$p_x\big(\cos u + c(y)\big) + p_y \sin u = 0.$$

Avec $u = \arg(p_x + ip_y)$, on a toujours :

$$p_x \cos u + p_y \sin u = p_x^2 + p_y^2,$$

d'où :

$$p_x^2 + p_y^2 + p_x c(y) = 0.$$

Comme $\sqrt{p_x^2 + p_y^2} > 0$ (sauf si $p_x = p_y = 0$, ce qui est exclu), cette équation implique $p_x < 0$. On en déduit :

$$\cos u = \frac{p_x}{\sqrt{p_x^2 + p_y^2}} = \frac{p_x}{-p_x c(y)} = -\frac{1}{c(y)}.$$

Cette expression est identique à celle obtenue dans la partie A pour la minimisation du déport.

Explication. Dans le cas anormal, la minimisation du temps n'est pas active ($p^0 = 0$), et la solution optimale coïncide avec celle du problème de minimisation du déport. Le contrôle optimal est déterminé par la nécessité de compenser le courant fort $c(y)$.

1.3.3 2.1 Condition de transversalité

Énoncé. On considère un problème de commande optimale à temps final fixé dont les conditions terminales sont (état de dimension 2)

$$x_1(t_f)\, x_2(t_f) = 1.$$

Donner la condition de transversalité correspondante.

Correction. La contrainte finale s'écrit sous la forme $h(x) = 0$ avec $h(x) = x_1x_2 - 1$. La différentielle de h est donnée par :

$$h'(x) = [x_2 \quad x_1],$$

qui est non nulle sur l'ensemble $\{h = 0\}$ (puisque x_1 et x_2 ne peuvent pas être simultanément nuls si $x_1x_2 = 1$).
La condition de transversalité impose que le vecteur adjoint final $p(t_f)$ soit orthogonal au noyau de $h'(x(t_f))$, c'est-à-dire que $p(t_f)$ soit colinéaire au gradient $\nabla h(x(t_f))$. Ainsi, il existe $\lambda \in \mathbb{R}$ tel que :

$$p(t_f) = \lambda \nabla h(x(t_f)) = \lambda\big(x_2(t_f), x_1(t_f)\big).$$

En éliminant λ, on obtient la condition :

$$p_1(t_f)\, x_1(t_f) - p_2(t_f)\, x_2(t_f) = 0.$$

Explication. La condition de transversalité exprime l'orthogonalité du vecteur adjoint final au noyau de la différentielle de la contrainte. Dans ce cas, elle se réduit à une relation algébrique entre les composantes de l'état final et des adjoints finaux.

1.3.4 2.2 Code de tir

Énoncé. En déduire comment compléter le code de tir ci-dessous :

```
function shoot(p0)
    xf, pf = f(t0, x0, p0, tf)
    s = [xf[1] * xf[2] - 1, pf[1] * xf[1] - pf[2] * xf[2]] # À COMPLÉTER
    return s
end
```

Correction. Le code est déjà complété correctement. Les deux composantes du vecteur s correspondent :

— la première à la contrainte finale sur l'état : $x_1(t_f)\,x_2(t_f) - 1 = 0$;

— la seconde à la condition de transversalité : $p_1(t_f)\,x_1(t_f) - p_2(t_f)\,x_2(t_f) = 0$.

Explication. La fonction `shoot` retourne un vecteur dont les composantes doivent être annulées par l'algorithme de tir. Ces conditions assurent que la solution respecte à la fois la contrainte finale sur l'état et les conditions nécessaires d'optimalité.

1.3.5 2.3 Vecteur `ts` dans MPC

Énoncé. Dans le code MPC ci-dessous, à quoi correspond le vecteur `ts` ?

```
while true
    w = drift(x1, y1)
    us, s = solve(x1, y1, 1, xf, yf, f, w, P, print_level=0)
    ts = [t1+s[1], t1+s[1]+s[2]]
    tf = t1+s[1]+s[2]+s[3]
    if (t1+t < tf)
        t2 = t1+t
    else
        t2 = tf
        println("t2=tf: ", t2)
    end
    sol = trajectory((t1, t2), x1, y1, 1, us, ts, drift)
    ...
end
```

Correction. Le vecteur `ts` correspond aux instants de commutation entre les différents arcs de la trajectoire optimale. Plus précisément, il contient les temps absolus auxquels le contrôle change de structure.

Explication. Dans cet algorithme de contrôle prédictif (MPC), on recalcule périodiquement la trajectoire optimale. Les valeurs `s` représentent les durées des différents arcs, donc `ts` donne les instants absolus où le contrôle commute. Cela permet de reconstruire la loi de commande optimale par morceaux.

1.3.6 2.4 Renforcement dans Hexapawn

Énoncé. On considère une partie d'Hexapawn pendant laquelle la machine vient de jouer le coup ci-dessous :

```
[ 2 2 0        [ 2 0 0
  1 0 1   ->     2 0 1
  0 0 1 ]        0 0 1 ]
```

La liste de coups de la machine associée à l'état précédent (= avant son dernier coup) était

```
{ [ 1 2   , [ 1 2   , [ 1 2
    2 1 ]     2 2 ]     2 3 ] }
```

Comment cette liste doit-elle être mise à jour par renforcement ?

Correction. Le coup joué par la machine est perdant (elle a perdu un pion sans compensation). Dans l'apprentissage par renforcement, on doit supprimer les coups menant à des positions défavorables. La liste mise à jour devient :

```
{ [ 1 2   , [ 1 2
    2 1 ]     2 3 ] }
```

Explication. L'algorithme de renforcement apprend en éliminant les actions qui conduisent à des états perdants. Ici, le coup `[1 2; 2 2]` a mené à une position désavantageuse, donc il est retiré de la liste des coups possibles pour cet état.