Chapitre 8

COMMANDE DU VÉHICULE REMI

INTRODUCTION.

Comme nous l'avons annoncé dans l'introduction, nous avons choisi comme mode de commande un asservissement latéral sur trajectoire, ou " path-following ", avec asservissement longitudinal indépendant. Justifions rapidement le choix de ce mode de commande. On l'oppose en général à l'asservissement sur la posture d'un robot de consigne virtuel, ou " path-tracking ", mode qui demande d'agir de manière coordonnée sur la vitesse et sur la direction du robot. D'un point de vue théorique, les deux modes de commande sont équivalents, au sens où l'on peut garantir la stabilisation du robot sur la trajectoire ou sur la posture de consigne, tant que la vitesse de consigne ne devient pas asymptotiquement nulle. Néanmoins, il est beaucoup plus simple de concevoir des asservissements latéral et longitudinal indépendants, qu'un asservissement sur une posture en mouvement le long de la trajectoire de consigne à la vitesse imposée, ou robot virtuel. Un autre argument en faveur de ce choix est que le suivi de la trajectoire est évidemment impératif, et prime devant le suivi de la vitesse. Or, dans le cas de l'asservissement sur posture, une perturbation provoquant un retard du processus sur le robot virtuel se traduirait inévitablement par un " raccourci " inacceptable par rapport à la trajectoire de celui-ci.

I. ASSERVISSEMENT LATÉRAL.

Enjeux de l'asservissement latéral.

Le problème de l'asservissement latéral a été largement étudié, en particulier par Claude Samson [SAM90] [SAM91a, b, &c] [SAM92]. L'auteur a établi qu'il est possible de stabiliser sur une trajectoire de consigne un robot mobile de type tricycle par retour d'état statique continu, tant que la vitesse du robot ne tend pas vers zéro, et donne dans [SAM92] un exemple d'une telle commande. Cependant, les lois de commande mises au point sont conçues pour des systèmes idéaux, c'est-à-dire purement cinématiques, comme le tricycle présenté au chapitre 7 (où v et β sont les entrées de commande). Elles ne peuvent donc être utilisées telles quelles pour la commande de systèmes réels, dont les actionneurs ont une dynamique non négligeable. Ainsi, une commande non linéaire par retour d'état statique issue de [SAM92] a été testée sur REMI sans qu'il y ait d'amélioration sensible par rapport à une commande linéaire, car elle tient compte de la non-linéarité (négligeable) de la cinématique, mais non de celle des actionneurs [MIN93a]. L'enjeu de la présente étude réside dans la mise en œuvre d'une loi de commande stabilisante, prenant en considération les non-linéarités du

véhicule réel, en effectuant l'apprentissage du correcteur latéral à l'aide du modèle de simulation latéral neuronal présenté au chapitre 7. Nous verrons plus loin que le problème terminal (qui se pose lorsque la vitesse de consigne s'annule) est résolu dans la pratique. En effet, que ce soit par une stratégie de path-following ou par une stratégie de path-tracking, la stabilisation sur une posture n'est effective à l'aide d'un retour d'état statique que si celui-ci est instationnaire, ou bien discontinu. Une solution instationnaire est proposée dans [SAM91c] [SAM92], une solution discontinue l'est dans [CAN91] ou [SOR92]. Une solution hybride (discontinue, et instationnaire au voisinage de la posture de consigne) est donnée dans [POM92].

Position du problème.

Soit une trajectoire de consigne (C) quelconque, respectant toutefois les contraintes non holonomes[1] du véhicule, en particulier la courbure maximale admissible de la trajectoire. Le problème de l'asservissement sur trajectoire consiste à déterminer une loi de commande pour l'angle du volant afin d'imposer au véhicule (i) de rejoindre et de suivre la trajectoire de consigne, et (ii) d'avoir son cap tangent à cette trajectoire quand il la suit. Nous avons choisi de réaliser l'apprentissage d'un régulateur stabilisant le véhicule sur une trajectoire rectiligne. Cette trajectoire rectiligne est ensuite choisie comme une tangente à la trajectoire réelle de consigne en un point appelé *point-cible*. Cette approche, fondée sur l'asservissement de la posture du véhicule, a été développée dans [RIV93a] et [RIV93b] (" posture-based approach "). Elle y est comparée à une approche fondée sur l'asservissement du cap du véhicule, respectant la séparation traditionnelle guidage-pilotage (" heading-based approach "), qui s'est révélée moins performante. Nous présentons maintenant l'apprentissage du régulateur neuronal correspondant à l'asservissement sur trajectoire ; nous considérons ensuite le choix du point-cible.

---

[1] Une *contrainte non holonome de type égalité* s'exprime par une contrainte scalaire non intégrable sur les paramètres de configuration (x, y et ψ pour la cinématique d'un robot se déplaçant dans un plan) et leurs dérivées ; elle n'affecte pas la dimension de l'espace des configurations, mais réduit la dimension de l'espace des vitesses. Par exemple, un véhicule tel que REMI peut adopter n'importe quelle configuration puisqu'il est commandable, mais ne peut se mouvoir parallèlement à l'axe de ses roues arrière. En supposant qu'il est parfaitement décrit par le modèle tricycle, ceci est dû à la contrainte non holonome de roulement pur qui s'écrit : $- x' \sin\psi + y' \cos\psi = 0$, et qui limite l'espace des vitesses au plan ainsi défini.

Une *contrainte non holonome de type inégalité* limite la vitesse à une partie de l'espace des vitesses. Par exemple, la limitation du rayon de courbure d'un chariot par la saturation en amplitude de l'angle de la roue restreint l'ensemble des vitesses possibles à un cône dans l'espace des vitesses [LAT91].

I.1. APPRENTISSAGE D'UN RÉGULATEUR NEURONAL DE STABILISATION SUR TRAJECTOIRE.

La trajectoire de consigne est la droite du plan d'équation y=0 (l'axe des abscisses), orientée vers les x positifs. Le but de l'asservissement est de stabiliser le véhicule sur cette trajectoire, c'est-à-dire de réguler les variables d'état ψ_p, le cap du véhicule, et y_p, son ordonnée. Puisque les consignes sont nulles, ces variables représentent aussi l'erreur de cap et l'erreur latérale. Nous traitons ce problème de régulation par la méthode de commande présentée au chapitre 5 §I, qui est donc une extension de la régulation LQ à horizon infini à des modèles non linéaires.

Modèle de simulation.

Les variables d'état du modèle sont le cap ψ, l'angle du volant β, l'abscisse x et l'ordonnée y du point de commande. La commande de l'angle du volant est notée α. Nous rappelons le modèle de simulation obtenu à l'issue de l'identification latérale :

$$\begin{cases} \beta(k) = sat_{\beta_{max}}\left(\beta(k\text{-}1) + sat_{\Delta\beta_{max}}\left(a_{lat}\left(\alpha(k\text{-}1) - \beta(k\text{-}1)\right)\right)\right) \\ \psi(k) = \psi(k\text{-}1) + \frac{\Delta T}{L}\, v(k\text{-}1)\ \psi_{RN}^{ModLat}\left(\beta(k\text{-}1)\right) \\ x(k) = x(k\text{-}1) + v(k\text{-}1)\ \Delta T\ cos\ \left(\psi(k\text{-}1)\right) \\ y(k) = y(k\text{-}1) + v(k\text{-}1)\ \ \Delta T\ \ sin\ \left(\psi(k\text{-}1)\right) \end{cases}$$

La vitesse v est considérée comme une perturbation mesurée. Le modèle et ses variables d'état sont représentés sur la figure 1. On note z=[β ψ y] les variables d'état d'intérêt pour la régulation.

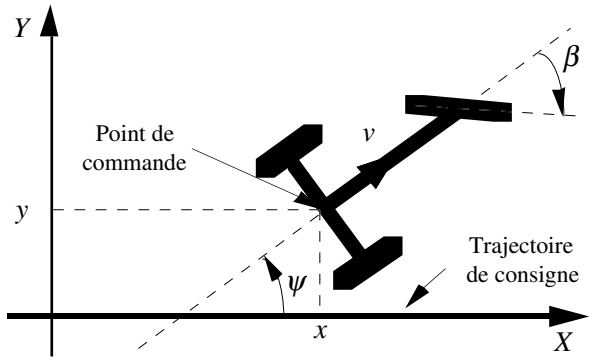

Figure 1.
Stabilisation sur la trajectoire de consigne y=0.

Régulateur neuronal.

Le réseau de neurones correcteur réalise un retour d'état statique, aussi fonction de la perturbation de vitesse :

$$\alpha(k) = \psi_{RN}\left(\beta(k),\ \psi(k),\ y(k),\ v(k);\ C\right)$$

Il n'y a pas de retour sur l'abscisse x, puisque l'asservissement longitudinal est indépendant. La non-linéarité du neurone de sortie est choisie sigmoïdale, à valeurs dans l'intervalle [$-\beta_{max}$; β_{max}], où β_{max} est l'amplitude maximale de l'angle du volant.

Séquences d'apprentissage.

Les séquences d'apprentissage sont constituées de N_Z trajectoires de N pas d'échantillonnage correspondant à une initialisation aléatoire (distribution uniforme) de l'état z(0) dans les intervalles [$-\beta_{max}$; β_{max}] pour l'angle du volant β (rd), [$-\pi$; π] pour le cap ψ (rd), et [-10; 10] pour l'ordonnée y (m) (l'abscisse initiale est toujours prise nulle, puisqu'elle n'intervient pas dans le problème). La vitesse est constante sur chaque trajectoire, initialisée aléatoirement dans l'intervalle [0; 20] (m/s).

Fonction de coût.

Il n'est pas nécessaire de faire intervenir la commande α dans la fonction de coût, puisque la sortie du régulateur est limitée (par une sigmoïde) :

$$J = \frac{1}{N_Z} \sum_{z(0)\in Z} \left(\sum_{k=1}^{N} z^T(k)\ W\ z(k) \right) = \frac{1}{N_Z} \sum_{z(0)\in Z} \left(\sum_{k=1}^{N} w_\beta\, \beta(k)^2 + w_\psi\, \psi(k)^2 + w_y\, y(k)^2 \right)$$

Faire intervenir l'angle du volant β reviendrait à pondérer la commande : nous avons donc choisi w_β=0. Pour déterminer la meilleure pondération du cap ψ et de l'ordonnée y, nous avons procédé de manière itérative, en imposant w_y=1 et en faisant varier w_ψ. Comme nous le verrons (figure 3), ce choix a une influence très importante sur l'allure des trajectoires.

Algorithme d'apprentissage.

L'algorithme est nécessairement semi-dirigé (cf chapitre 5 §I.3.1). Le système d'apprentissage est représenté sur la figure 2.

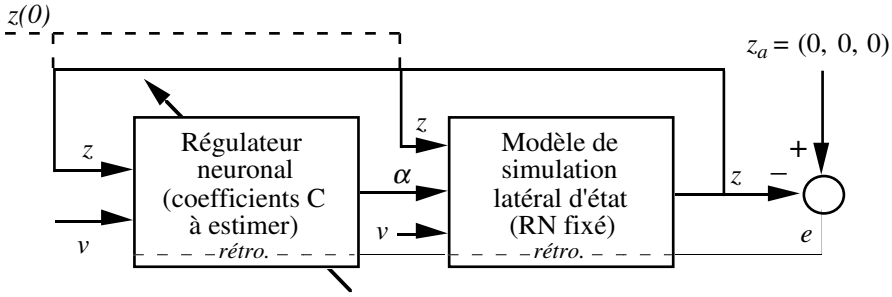

Figure 2.
Système d'apprentissage pour l'asservissement sur trajectoire rectiligne (y=0).

Plusieurs apprentissages ont donc été effectués pour déterminer les coefficients de pondération. La figure 3 présente le comportement du modèle commandé résultant d'apprentissages du régulateur pour trois valeurs différentes de w_ψ : 1, 10 et 50. Ces trajectoires de test, représentatives de la dynamique de ralliement, ont été obtenues en initialisant le modèle neuronal du véhicule sur la trajectoire avec un cap variant entre $-\pi$ et $+\pi$ et un angle du volant de 0 rd, à une vitesse constante de 10 m/s (la forme des trajectoires varie assez peu en fonction de la vitesse). Ces simulations nous ont conduit au choix de la pondération suivante : $w_\beta = 0$; $w_\psi = 10$; $w_y = 1$. La performance ne s'améliore plus à partir de six neurones cachés.

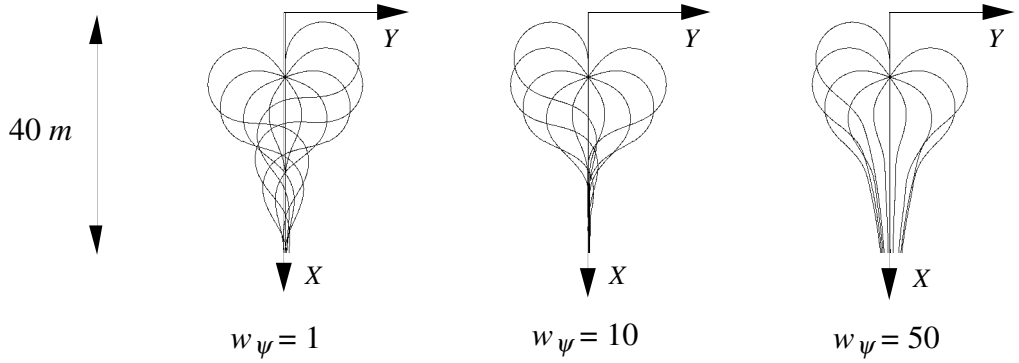

Figure 3.
Influence de la pondération des variables d'état dans la fonction de coût sur la dynamique de ralliement de la trajectoire.

I.2. CHOIX D'UN POINT-CIBLE.

Comme cela a été dit, le point-cible est le point de la trajectoire (C) choisi à chaque instant pour déterminer la tangente à la trajectoire sur laquelle s'asservir. Si l'on choisit comme point-cible le point de (C) le plus proche du véhicule, il faut tenir compte de la courbure de la trajectoire en ce point. La solution la plus simple consiste à ajouter un terme de boucle ouverte de la forme[2] :

$$\alpha_c(k) = L\ \kappa(k)$$

où κ(k) est la courbure de (C) au point-cible défini à l'instant k. Cette solution est utilisée par exemple par [RIN93] [VDM93]. Elle se révèle cependant peu performante dans le cas de courbures importantes de la trajectoire (nous l'avons expérimentée). En effet, la relation ci-dessus suppose que le robot se trouve déjà sur la trajectoire, donc dès que ce n'est plus le cas, typiquement dans le cas de virages importants, la loi de commande est inadéquate[3].

Sans remettre en question l'apprentissage sur trajectoire de courbure nulle effectué (il est possible de réaliser un apprentissage sur des cercles de courbure variant entre 0 et κ_{max}, en fournissant la courbure en entrée du réseau), il est un autre moyen de tenir compte de la courbure, cette fois de façon implicite, tout en évitant les inconvénients de l'approche fondée sur la courbure. Ce moyen consiste à choisir le point-cible non pas au plus près sur la trajectoire, mais à une certaine distance D en avant du véhicule, afin d'anticiper les virages éventuels. Ce choix du point-cible et de la tangente à la trajectoire en ce point est représenté sur la figure 4.

La distance D est choisie proportionnelle à la vitesse : $D = F_D\ v_p$. Le facteur de proportionnalité F_D est fixé en simulation, en faisant évoluer le véhicule simulé sur des trajectoires réelles cartographiées. Les simulations permettent de fixer $D = v$ (soit F_D=1), où v est la vitesse du modèle.

---

[2] Cette relation est obtenue à partir de la variation de cap du modèle linéarisé. Si l'on suppose que ce modèle suit la trajectoire, on a : dψ/dt = ds/dt 1/L α, soit α = L dψ/ds = L κ (où s est l'abscisse curviligne le long de la trajectoire).

[3] Pour éviter cet inconvénient, la loi de commande donnée dans [SAM92] comporte un terme sur la courbure qui s'écrit : $\dot{\alpha}_c = v\ \kappa\ cos\left(\psi\right) / \left(1 - \kappa\ y\right)$. Ce terme varie donc en fonction de l'erreur latérale.

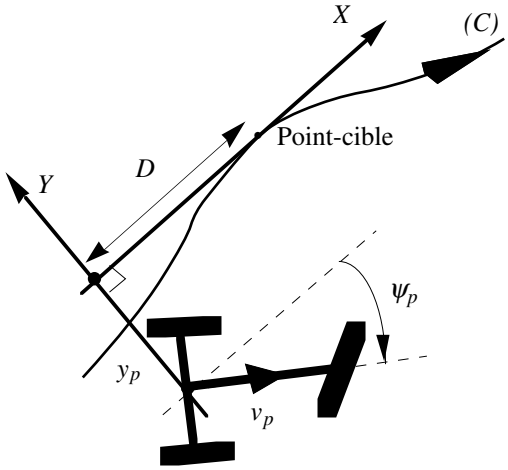

Figure 4.
Choix du point-cible et détermination de la tangente sur laquelle asservir le point de commande.

Justification du choix $D = F_D\, v_p$.

Le choix d'une distance au point-cible proportionnelle à la vitesse mérite une justification. Nous allons montrer qu'aux petites erreurs (erreur latérale y et erreur de cap ψ), l'application d'un retour d'état conçu pour rallier une trajectoire rectiligne pour le choix du point-cible comme point le plus proche de (C) avec addition d'un terme proportionnel à la courbure (approche " courbure "), est équivalente à l'application du même retour d'état pour le choix d'un point-cible évoluant à une distance D proportionnelle à la vitesse en avant du point de commande (approche " point-cible ").

Pour de petites erreurs, on peut raisonner avec le modèle linéarisé suivant (nous nous plaçons en temps continu pour simplifier la présentation) :

$$\begin{cases} \dot{\psi} = \frac{v}{L}\, tg\,(\alpha) \approx \frac{v}{L}\,\alpha \\ \dot{y} = v\ \ sin\,(\psi) \approx v\ \ \psi \end{cases}$$

On montre facilement que pour imposer une dynamique de ralliement de la trajectoire y=0 indépendamment de la vitesse, il faut utiliser un retour d'état de la forme :

$$\alpha = K_y\, y + K_\psi\, \psi \qquad avec \qquad \begin{cases} K_y = \frac{k_y}{v^2} & k_y < 0 \\ K_\psi = \frac{k_\psi}{v} & k_\psi < 0 \end{cases}$$

(Le réseau de neurones trouve donc nécessairement une solution de ce type aux petites erreurs.)

Notons α_{cou} et α_{pc} les commandes utilisées respectivement dans les approches courbure et point-cible. Ces commandes ont pour expressions :

$$\begin{cases} \alpha_{cou} = K_y\, y_{cou} + K_\psi\, \psi_{cou} + L\, \kappa \\ \alpha_{pc} = K_y\, y_{pc} + K_\psi\, \psi_{pc} \end{cases}$$

où y_{cou} et ψ_{cou} sont les variables d'état de l'approche courbure, et y_{pc} et ψ_{pc} les variables d'état de l'approche point-cible ; κ est la courbure de la trajectoire entre le point le plus proche et le point-cible, supposée constante. Ces variables sont représentées sur la figure 5.

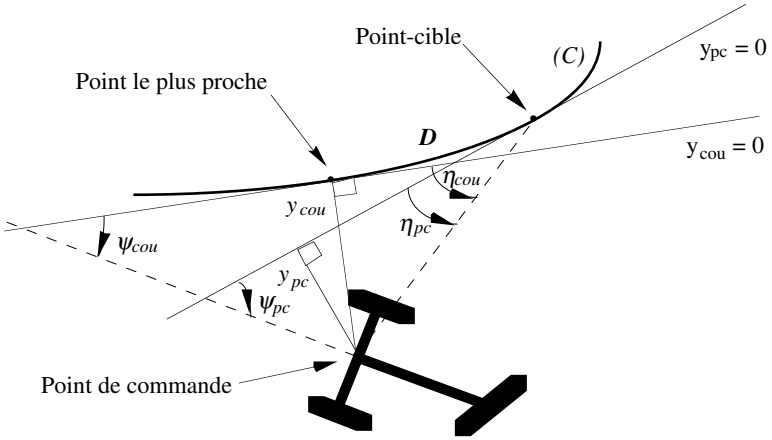

Figure 5.
Comparaison des approches fondées sur la courbure et sur le point-cible.

Aux petits angles, on a (en utilisant la relation angulaire entre les angles η_{cou} et η_{pc}) :

$$\begin{cases} y_{pc} = y_{cou} - \dfrac{D^2}{2}\,\kappa \\ \psi_{pc} = \psi_{cou} + D\,\kappa \end{cases}$$

Récrivons α_{pc} :

$$\alpha_{pc} = K_y\, y_{cou} + K_\psi\, y_{cou} + \left(K_\psi\, D - K_y\, \frac{D^2}{2}\right)\,\kappa$$

Comme K_y et K_ψ sont inversement proportionnels à la vitesse et à son carré respectivement, cette expression de α_{pc} est égale à celle de α_{cou} si l'on prend $D = F_D\, v$, avec F_D tel que :

$$k_\psi f - k_y \frac{f^2}{2} = L$$

Ceci est toujours possible, l'équation du second degré :

$$f^2 - 2\frac{k_\psi}{k_y} f + 2\frac{L}{k_y} = 0$$

ayant toujours deux racines réelles dont une positive, puisque $k_y<0$.

Nous avons donc prouvé que, aux petites erreurs, et pour un régulateur donné imposant une dynamique de ralliement indépendante de la vitesse, une stratégie utilisant un point-cible situé à une distance proportionnelle à la vitesse du point le plus proche de (C) est fondée. Aux grandes erreurs, il est intuitif que cette solution ne possède pas les défauts de l'approche fondée sur la courbure, mais il est moins simple de calculer, comme nous venons de le faire, la loi de variation de la distance au point-cible.

I.3. RÉSULTATS EXPÉRIMENTAUX.

Nous montrons ici quelques résultats expérimentaux obtenus avec le véhicule sur terrain sec non accidenté, puis en dévers. Des résultats obtenus en tout terrain sont présentés au §III de ce chapitre (avec l'asservissement de vitesse correspondant).

Système de commande.

Il s'agit d'un système de commande par simple bouclage. Le régulateur latéral obtenu en fin d'apprentissage, que nous appelons " RégLat " est utilisé avec les arguments :

$$\alpha(k) = \psi_{RN}^{RégLat}\left(\beta_p(k),\ \psi_p(k),\ y_p(k),\ v_p(k)\right)$$

où y_p et ψ_p sont déterminés en fonction du point-cible sélectionné (voir figure 4). Le système de commande de l'asservissement latéral est représenté sur la figure 6.

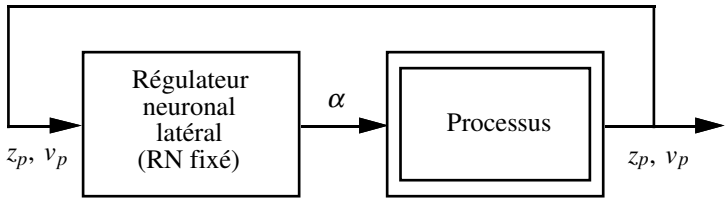

Figure 6.
Système d'asservissement latéral.

Asservissement latéral sur trajectoire.

La première expérience a pour but l'asservissement du véhicule sur une trajectoire de consigne d'environ 600 m de long (il s'agit du tour du centre de SAGEM Éragny). La courbure maximale de la trajectoire est de 0,1 m^{-1}. Cette trajectoire est représentée sur la figure 7.

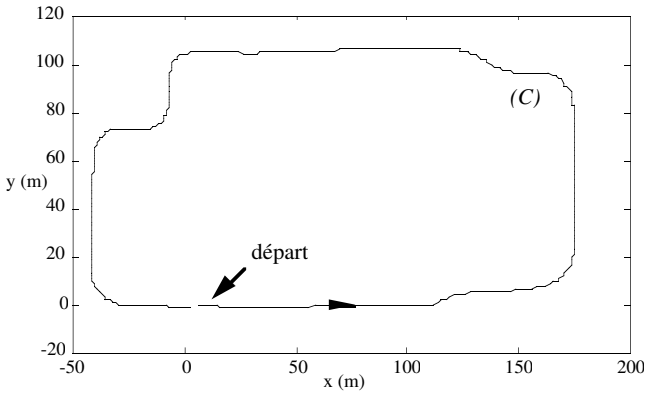

Figure 7.
Trajectoire de consigne.

Les résultats de l'asservissement sont rassemblés sur la figure 8. La vitesse, imposée par un conducteur, est assez importante compte tenu des virages serrés de la trajectoire (voir figure 8d). La figure 8c permet de visualiser la saturation en vitesse de l'angle du volant, qui se manifeste en particulier dans les virages situés en haut à gauche sur la trajectoire de consigne. L'erreur latérale indiquée est la distance du véhicule au point le plus proche de la trajectoire de consigne, et l'erreur de cap, l'angle de l'axe du véhicule par rapport à la tangente à la trajectoire en ce point. Ces erreurs, qui correspondent aux erreurs y_{cou} et ψ_{cou} de la figure 5, définissent la performance, et ne sont pas celles qui sont fournies au réseau pendant le fonctionnement du système de commande, y_{pc} et ψ_{pc}. Les performances sont bonnes, les deux erreurs étant inférieures à 40 cm et 50 mrd respectivement.

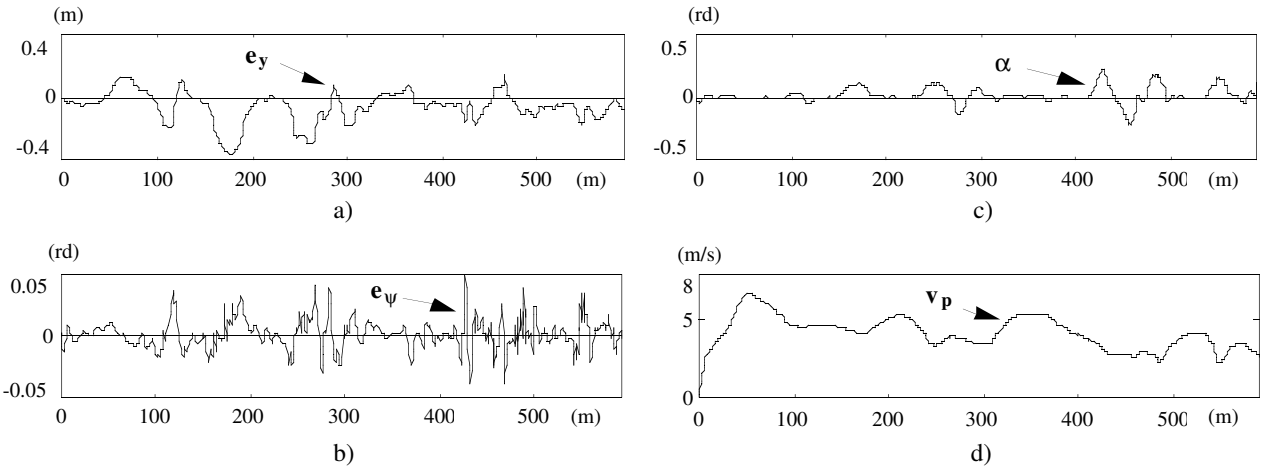

Figure 8.
Asservissement latéral sur trajectoire : a) erreur latérale, b) erreur de cap, c) commande du volant, d) vitesse, en fonction de l'abscisse curviligne le long de la trajectoire de consigne.

Ralliement de la trajectoire de consigne.

La deuxième expérience a été effectuée en vue de montrer la capacité du système de commande à rallier la trajectoire après que le véhicule en a été écarté de quelques mètres, c'est-à-dire hors du domaine des petites erreur de cap et erreur latérale. Nous présentons sur la figure 9 la trajectoire de consigne, et celle du point de commande du véhicule (milieu de l'essieu arrière).

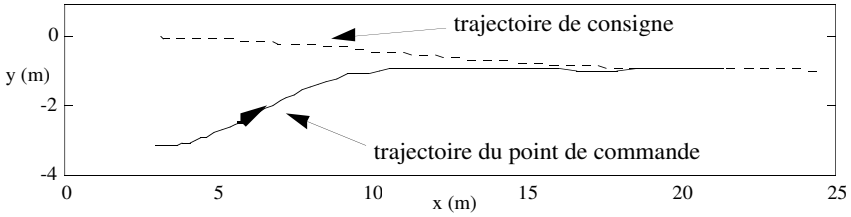

Figure 9.
Ralliement de la trajectoire de consigne.

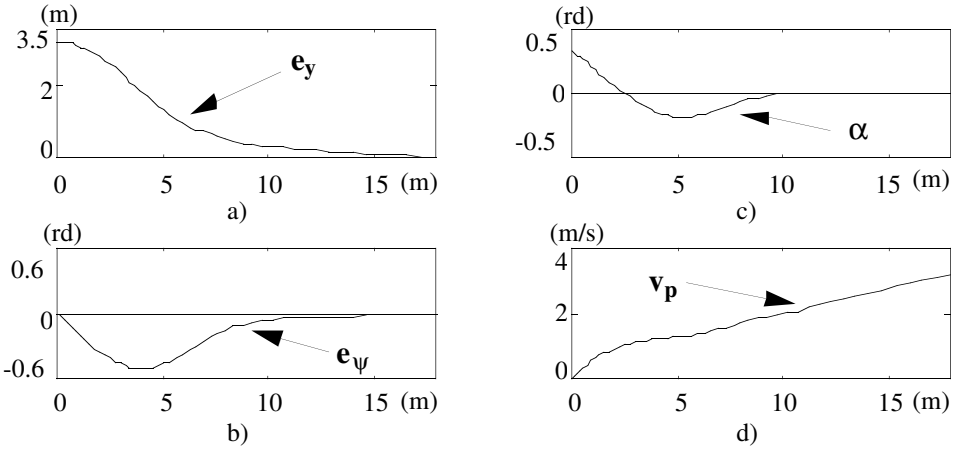

Figure 10.
Ralliement de la trajectoire de consigne : a) erreur latérale, b) erreur de cap, c) commande du volant, d) vitesse, en fonction de l'abscisse curviligne le long de la trajectoire de consigne.

L'évolution des variables d'intérêt le long de cette trajectoire est représentée sur la figure 10. Ici encore, la vitesse est imposée par un conducteur. À l'instant initial, le véhicule est maintenu à l'arrêt : le volant est alors complètement braqué au démarrage ($\alpha(0)=\beta_p(0)=0{,}5$ rd). Néanmoins, l'erreur latérale, initialement de 3,5 m, est annulée sans dépassement.

Asservissement latéral en dévers.

Le système de commande utilisé n'impose pas une erreur latérale statique nulle, puisque le correcteur ne contient pas d'intégrateur (explicite ou implicite). Il faut donc s'attendre à une erreur non nulle en cas de dévers important. Nous avons réalisé des essais sur des pistes en dévers calibrées sur les terrains d'essais de l'ETAS (Etablissement Technique d'AngerS). La trajectoire de consigne et celle du point de commande du véhicule sont représentées sur la figure 11.

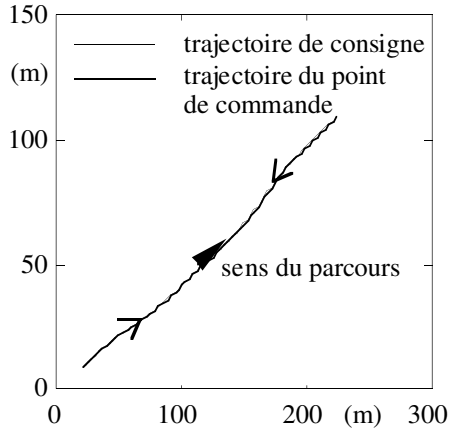

Figure 11.
Asservissement latéral en dévers (20%) à 15 km/h : trajectoires de consigne et du point de commande.

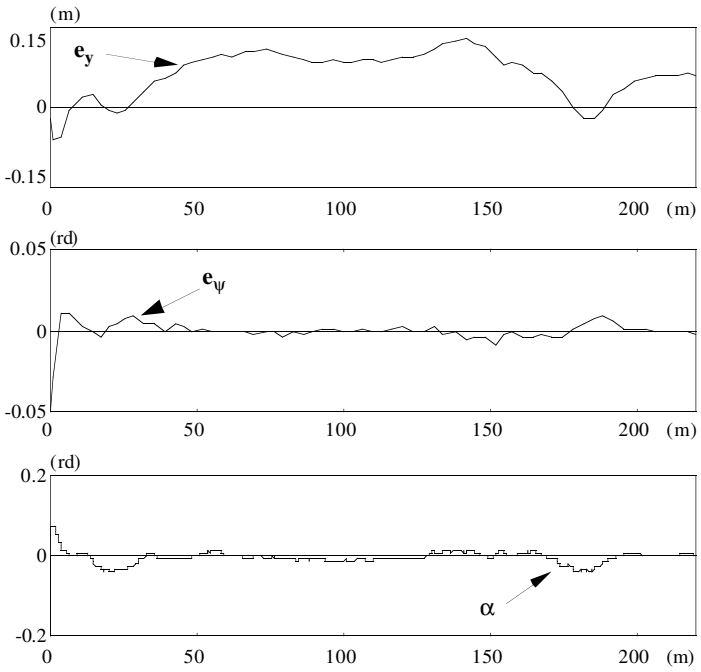

Figure 12.
Résultats d'asservissement latéral en dévers (20%) à la vitesse de 4,2 m/s (15 km/h).

Les résultats de la figure 12 correspondent à une piste en dévers de 20%, parcourue à une vitesse stabilisée de 15 km/h. L'erreur statique observée est de l'ordre de 10 cm, performance tout à fait satisfaisante. Les deux créneaux importants sur le volant (vers 25 et 175 m d'abscisse curviligne) correspondent aux endroits, indiqués par des chevrons sur la figure 11, où le véhicule aborde et quitte le dévers. À 6 km/h, toujours sur un dévers de 20%, l'erreur latérale n'est que de 4 cm environ ; à 6 km/h, sur un dévers de 30%, elle est de l'ordre de 7 cm [ETA94]. Il est préférable de se contenter de ces performances, très satisfaisantes, plutôt que d'introduire un terme intégral dans la loi de commande, qui risquerait de diminuer la stabilité du système.

II. ASSERVISSEMENT LONGITUDINAL.

Enjeux de l'asservissement longitudinal.

L'asservissement longitudinal de la vitesse s'effectue par l'action sur le papillon des gaz, sur la pression dans le circuit de freinage, et également sur le sélecteur des vitesses. Il n'existe pas de modèle universel de la dynamique longitudinale de robots mobiles à roues, car celle-ci dépend étroitement de leur mode de propulsion et de leurs actionneurs particuliers. Par exemple, on ne peut comparer la dynamique d'un véhicule électrique, dont un modèle standard est donné par exemple dans [SHL78], et celle d'un véhicule à moteur thermique tel que REMI. Plus encore que pour l'asservissement latéral, il est donc nécessaire de tenir compte de la dynamique particulière du véhicule. De plus, la dynamique de REMI est complexe et non linéaire, et l'expérience a montré qu'un simple PID était insuffisant pour une commande performante [MIN93b].

Position du problème.

Un profil de vitesse de consigne est associé à la trajectoire de consigne. Le problème de l'asservissement longitudinal est le problème de la poursuite de cette consigne de vitesse, problème que nous avons choisi de résoudre à l'aide d'un système de commande avec modèle interne (SCMI). Nous avons fait ce choix en raison de la moindre précision du modèle longitudinal dans certains domaines de fonctionnement (par comparaison avec le modèle latéral), et en fonction des résultats obtenus avec des systèmes de commande qui, comme on pouvait s'y attendre, se sont révélés moins performants : un système de commande par simple bouclage (SCSB) utilisant un correcteur-S et un modèle de ralliement, et un " PID " non linéaire (ou NID, voir chapitre 5 §II.3.3).

II.1. APPRENTISSAGE D'UN CORRECTEUR-S.

Le problème est donc celui de la poursuite et de la régulation de la vitesse du véhicule. On agit toujours séparément sur les freins et sur l'accélérateur, qui sont des commandes de même nature : le problème peut donc être considéré comme un problème à une seule entrée. Comme le modèle du processus est de type entrée-sortie de retard unité, nous avons vu que le SCMI le plus performant et

souple d'utilisation, dans ce cas, est le système utilisant un correcteur-S et un modèle de ralliement (chapitre 5 §II.4.2.1). Nous avons donc procédé à l'apprentissage d'un correcteur-S.

Modèle de référence d'apprentissage.

Le modèle de référence d'apprentissage, de sortie v_a, est un simple retard :

$$v_a(k+1) = v_r(k)$$

où v_r est la séquence de consigne.

Modèle de simulation.

La sortie du modèle est la vitesse v. Les entrées de commande du modèle sont l'angle papillon commandé θ, et la pression de freinage commandée π. Les entrées perturbatrices simulées sont d'une part le rapport des vitesses r, converti en la grandeur quantifiée ρ, et d'autre part la pente p. K est le facteur d'inertie qui est fonction du rapport engagé. Rappelons l'expression de ce modèle de simulation établi à partir du prédicteur identifié :

$$v(k) = \varphi_{RN}^{ModLon}\left(v(k\text{-}1),\, \theta(k\text{-}1),\, \pi(k\text{-}1),\, \rho(k\text{-}1)\right) - \frac{g\,\Delta T}{1 + K\left(\rho(k\text{-}1)\right)}\, sin\left(p(k\text{-}1)\right)$$

Correcteur-S.

Le réseau de neurones correcteur est non bouclé de la forme :

$$u(k) = \psi_{RN}\left(v_r(k), v(k),\, \rho(k),\, p(k);\, C\right)$$

où la commande u est appliquée à l'accélérateur si elle est positive, aux freins sinon :

$$\begin{cases} si\ u(k) \geq 0,\ \theta(k) = u(k)\ et\ \pi(k) = 0. \\ si\ u(k) < 0,\ \theta(k) = 0\ et\ \pi(k) = u(k). \end{cases}$$

Les valeurs autorisées de la commande u étant confinées dans l'intervalle [-1; 1], on choisit comme fonction d'activation du neurone de sortie une sigmoïde à valeurs dans [-1; 1].

Séquences d'apprentissage.

Les séquences d'apprentissage sont constituées de trajectoires à rapport des vitesses ρ et pente p constants. Ces trajectoires $\{v_r\}$ sont calculées par un modèle de référence à partir de créneaux de consigne $\{v_c\}$: ce modèle est le même quels que soient ρ et p ; il correspond à un filtre à temps continu du second ordre de pulsation ω_n=1 et d'amortissement ζ=1. Ce modèle est plus rapide que tous les modèles de ralliement qui seront utilisés au sein du SCMI. Il est noté :

$$E(q)\, v_r(k+1) = H(q)\, v_c(k)$$

où $\{v_c\}$ est une séquence de consigne constituée de 10 créneaux de durée 35 secondes. L'amplitude des créneaux de consigne dépend des valeurs de ρ et de p.

* Ainsi, pour une pente nulle, l'amplitude maximale v_{max} des créneaux est la suivante :

- en 1ère (ρ=0,00) : v_{max} = 10 m/s ;
- en 2nde (ρ=0,40) : v_{max} = 15 m/s ;
- en 3ème (ρ=0,64) : v_{max} = 20 m/s ;
- en 4ème (ρ=0,75) : v_{max} = 25 m/s.

La vitesse maximale choisie pour les trois premières vitesses est supérieure à celle que le modèle peut effectivement atteindre. En effet l'utilisation future du réseau dans un SCMI exige un apprentissage dans un domaine plus large que le domaine de fonctionnement réel du processus, même si le correcteur sature (voir chapitre 5 §II.4.1.5).

* Nous avons choisi de ne compenser la pente à l'aide du correcteur qu'en première : c'est en effet le seul rapport de vitesses qui sera utilisé pour gravir ou descendre des pentes importantes. Si, en fonctionnement, la première n'est pas imposée au moyen du levier des vitesses, ce qui ne sera le cas que pour des pentes faibles, le SCMI devra compenser. En première, on prend pour la pente p, pendant l'apprentissage, des valeurs comprises entre ±40%.

Fonction de coût.

La fonction de coût est la somme des carrés des erreurs de vitesse sur toutes les trajectoires :

$$J = \frac{1}{N}\sum_{k=1}^{N} e(k)^2 = \frac{1}{N}\sum_{k=1}^{N} \left(v_a(k) - v(k)\right)^2$$

où N est le nombre total d'instants de l'ensemble d'apprentissage.

Algorithme d'apprentissage.

L'algorithme d'apprentissage est semi-dirigé. Le système d'apprentissage est représenté sur la figure 13.

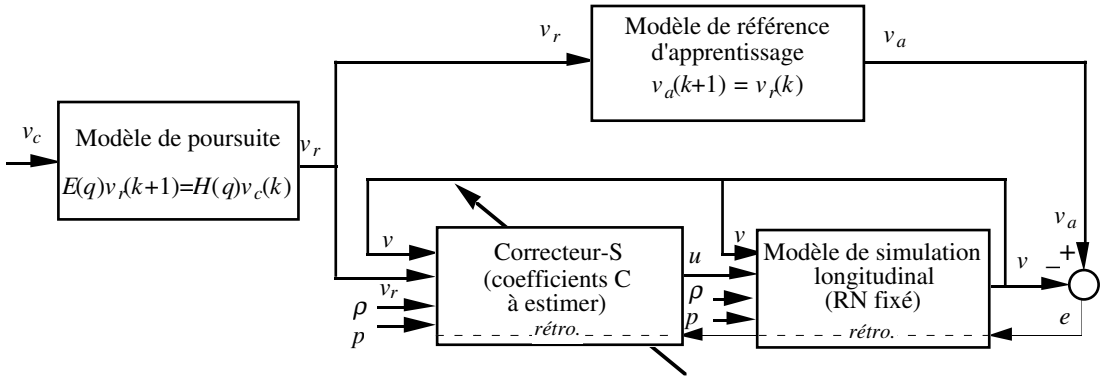

Figure 13.
Système d'apprentissage du correcteur-S.

Résultats d'apprentissage.

Nous allons maintenant présenter les résultats de cet apprentissage sur des séquences de test. Pour alléger la présentation, nous la restreignons à des simulations en pente nulle, la seule perturbation mesurée à compenser étant donc le rapport des vitesses.

Sur les figures 14a et 14b sont représentées les réponses du système d'apprentissage à des séquences de test dans le domaine de fonctionnement *souhaité pour le processus* : la consigne y est telle que le modèle peut l'atteindre en régime statique (la commande peut toutefois se saturer si l'accélération demandée est trop importante). Les limites de saturation de la commande à ±1 sont indiquées en traits fins pointillés sur les figures.

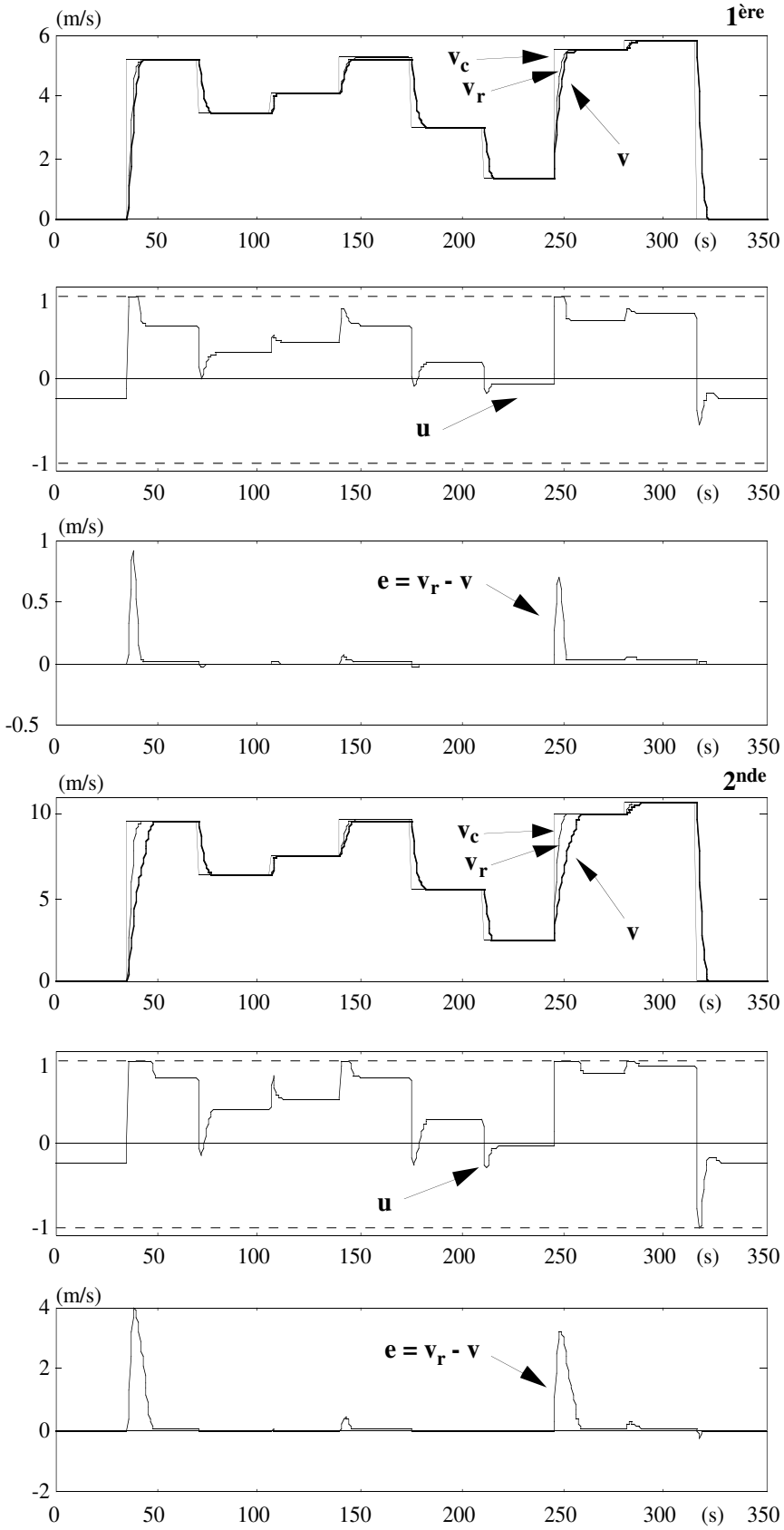

Figure 14a.
Test du correcteur-S en simple bouclage avec le modèle, en 1ère et en 2nde, dans le domaine de fonctionnement souhaité pour le processus, avec un modèle de poursuite rapide (ω_n=1).

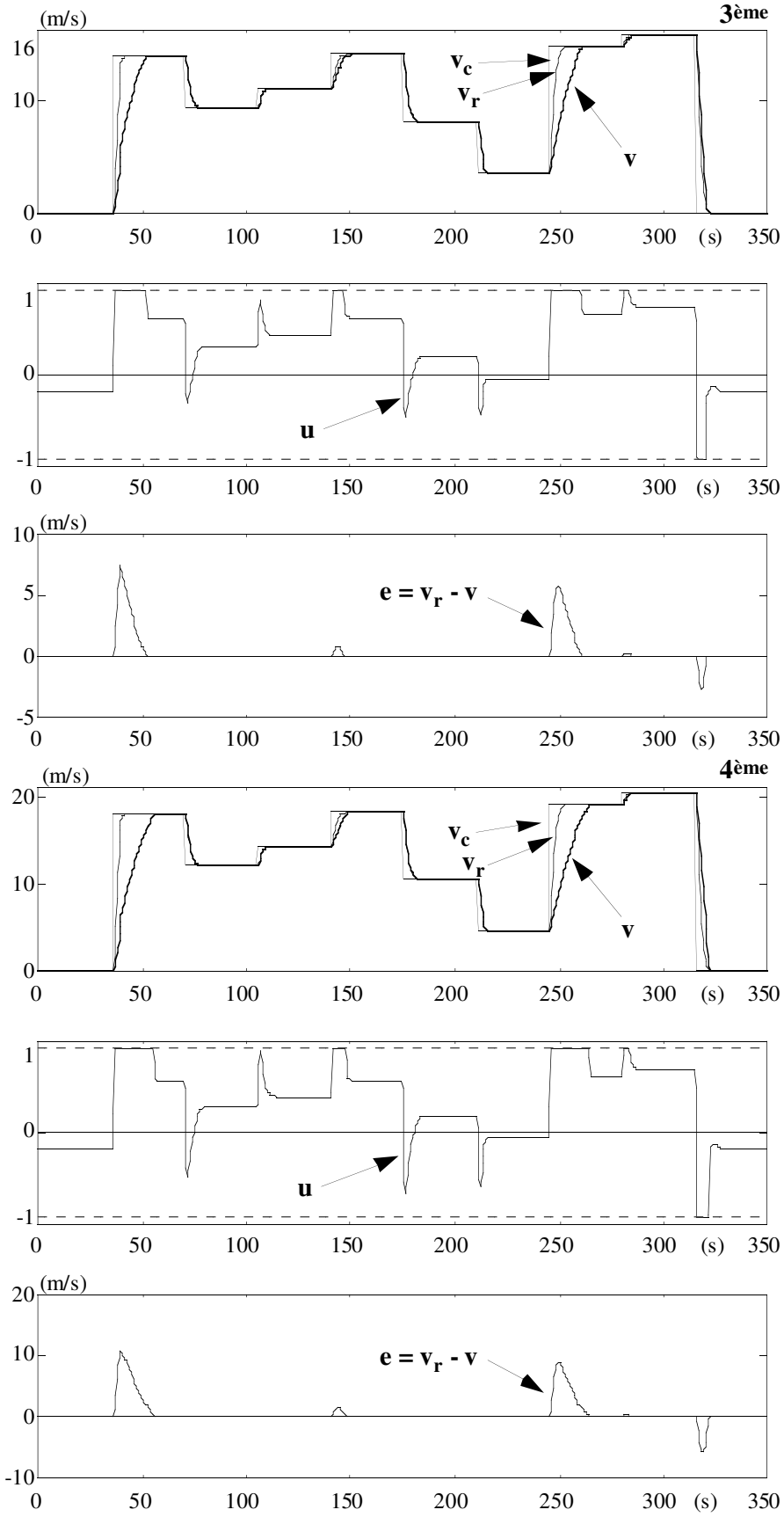

Figure 14b.
Test du correcteur-S en simple bouclage avec le modèle, en 3ème et en 4ème, dans le domaine de fonctionnement souhaité pour le processus, avec un modèle de poursuite rapide (ω_n=1).

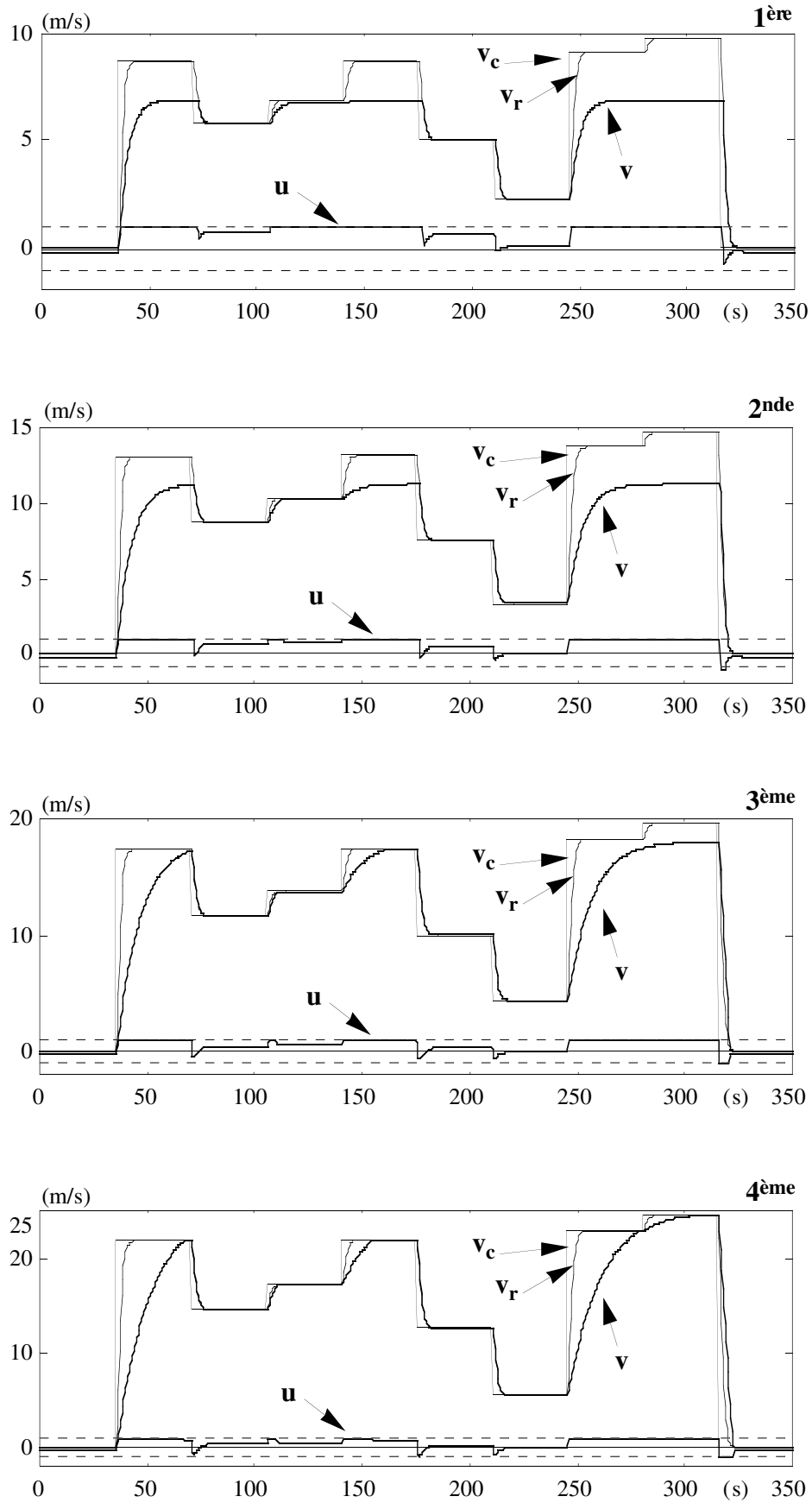

Figure 15.
Test du correcteur-S en simple bouclage avec le modèle, dans le domaine d'apprentissage entier du modèle, avec un modèle de poursuite rapide (ω_n=1) : le correcteur est presque toujours **saturé** (u=±1).

Sur la figure 15, nous avons rassemblé des résultats illustrant le comportement du modèle commandé dans tout le domaine d'apprentissage, qui est le domaine de fonctionnement *souhaité pour le modèle* : les consignes maximales y sont telles que, étant donnés la saturation en ±1 du correcteur et le gain statique du modèle, celui-ci ne peut les atteindre. Il est capital que le comportement du réseau correcteur soit correct pour de telles consignes, qui ne seraient pas demandées dans un SCSB, mais qui peuvent l'être dans un SCMI (car la consigne y est décalée de la valeur de l'écart de vitesse entre le processus et le modèle). Les courbes de la figure 15 montrent que le comportement du correcteur est effectivement correct dans tout le domaine d'apprentissage.

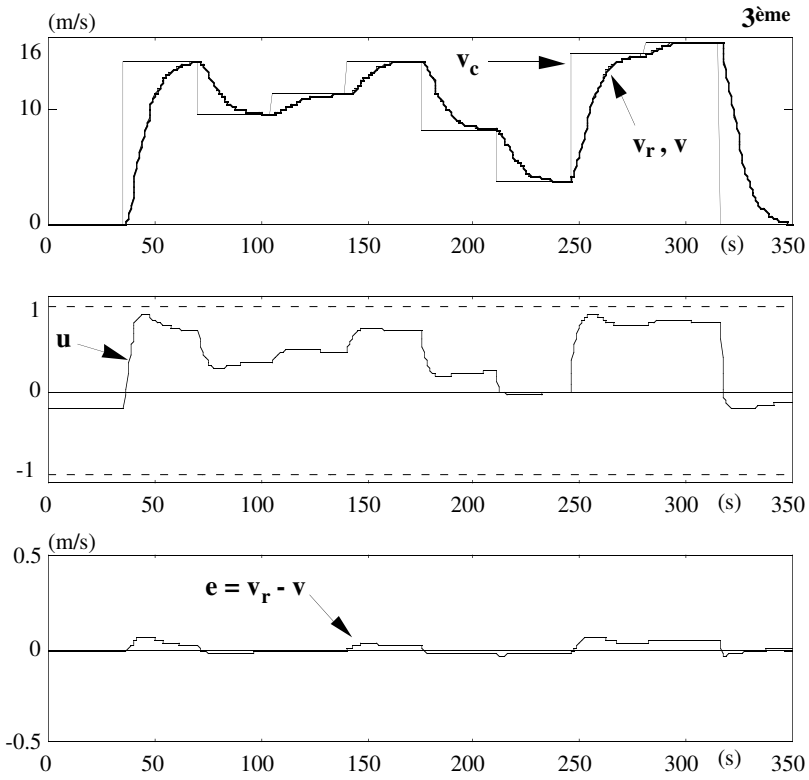

Figure 16.
Test du correcteur-S en simple bouclage avec le modèle, dans le domaine de fonctionnement souhaité pour le processus, en 3ème, avec un modèle de poursuite plus lent (ω_n=0,2).

La figure 16 enfin illustre le comportement du système d'apprentissage pour des séquences de consignes obtenues avec un modèle de référence plus lent (correspondant à la discrétisation d'un filtre continu de pulsation ω_n=0,2 et ξ=1) ; elle est donc représentative des performances que l'on peut obtenir pour des entrées de consigne moins rapides que des échelons : dans ces conditions, le correcteur-S est presque parfait (l'erreur de commande est inférieure à 5 10^{-2}, lorsque la commande ne se sature pas). Nous montrons les résultats obtenus en 3ème, dans le domaine de fonctionnement souhaité pour le processus. Ces résultats sont à comparer à ceux de la figure 14, en 3ème également.

Ces résultats sont obtenus à l'aide d'un correcteur possédant 5 neurones cachés complètement connectés. Le correcteur est donc plus complexe que le modèle (3 neurones cachés). Ceci est dû au

fait que la commande est appliquée, suivant son signe, à deux entrées distinctes du modèle, de caractéristiques très différentes (celle des freins et celle de l'accélérateur).

II.2. RÉSULTATS EXPÉRIMENTAUX (COMMANDE AVEC MODÈLE INTERNE).

Système de commande.

Comme nous l'avons annoncé, le correcteur-S est intégré dans un SCMI, avec un modèle de ralliement. Ce modèle de ralliement définit une dynamique identique à celle du modèle de poursuite utilisé pendant l'apprentissage pour calculer la séquence de consigne (ou éventuellement plus lente, en fonction de résultats d'essais avec le véhicule). L'expression de sa sortie v_r* est donnée par :

$$v_r^*(k+1) = \left(1 - E(q)\right) v(k+1) + H(q)\, v_c^*(k)$$

Le correcteur-S obtenu, que nous appellons ψ_{RN}^{CorLon}, a pour arguments :

$$u(k) = \psi_{RN}^{CorLon}\left(v_r^*(k+1),\, v(k),\, \rho_p(k),\, p_p(k)\right)$$

où ρ_p et p_p sont les estimations du rapport des vitesses du véhicule et de la pente du terrain.
Le modèle interne est bouclé, et défini par la même fonction que le modèle de simulation entrée-sortie de l'apprentissage, mais avec les arguments :

$$v(k) = \varphi_{RN}^{ModLon}\left(v(k\text{-}1),\, \theta(k\text{-}1),\, \pi(k\text{-}1),\, \rho_p(k\text{-}1)\right) - \frac{g\,\Delta T}{1 + K\left(\rho_p(k\text{-}1)\right)} \sin\left(p_p(k\text{-}1)\right)$$

Le modèle de ralliement correspond au filtre du second ordre avec ω_n=0,3 et ξ=1 (ce choix résulte d'expérimentations avec le véhicule). Le système de commande est représenté sur la figure 17.

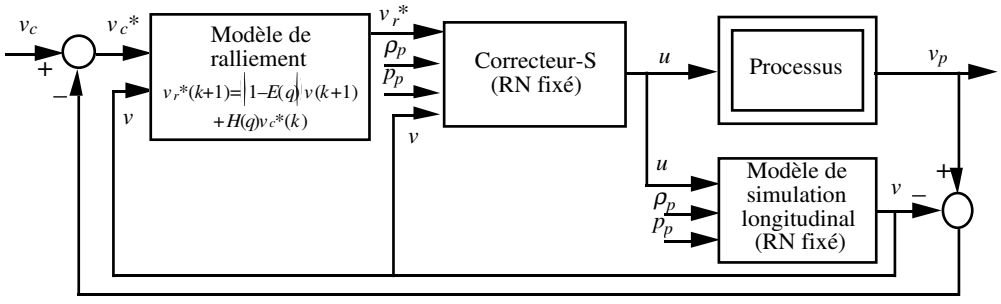

Figure 17.
SCMI avec correcteur-S et modèle de ralliement pour l'asservissement longitudinal.

Modalités d'expérimentation.

La figure 18 montre ainsi un exemple des résultats expérimentaux obtenus sur terrain plat (route). Avant de commenter ces résultats, nous donnons quelques détails sur les modalités de l'expérimentation.

- Comme pour l'asservissement latéral, l'asservissement longitudinal est ici mis en œuvre sans asservissement latéral, un conducteur assurant le pilotage de la direction du véhicule ;
- La position du levier des vitesses est telle que le véhicule peut passer tous les rapports, ce que l'on peut voir sur la figure 18 (la valeur de $\rho_p = 1 - 1/r_p$ atteint sa valeur maximale 0,75). La valeur du

rapport donnée en entrée du réseau de neurones ρ_p est quantifiée selon les quatre valeurs possibles du rapport (à partir de l'estimation faite en fonction de la vitesse et du régime moteur) ;

- Lorsque la vitesse de consigne est nulle, et que le véhicule est presque à l'arrêt, la commande maximale de freinage u=-1 est appliquée, par sécurité (instants 280 et 380 sur la figure 18).

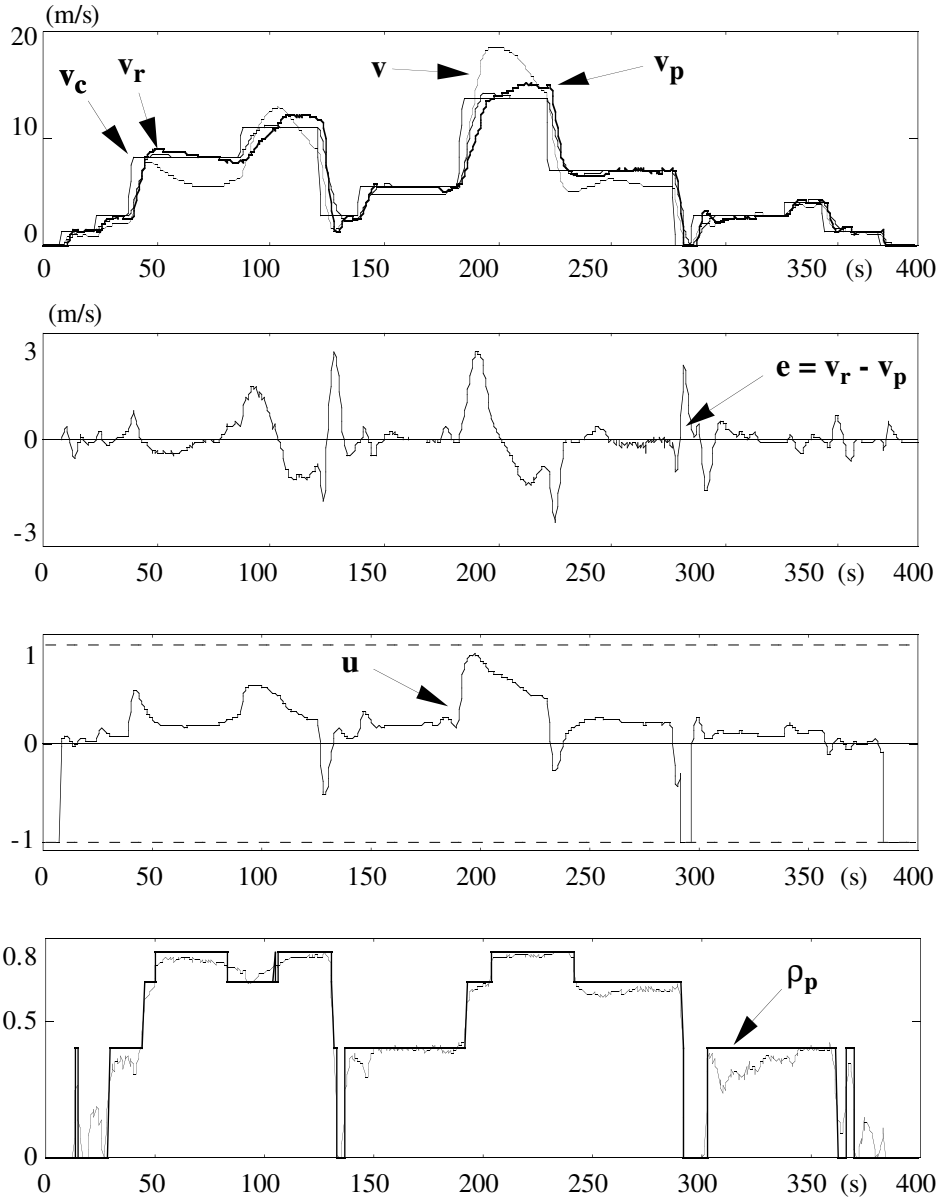

Figure 18.
Résultats expérimentaux obtenus avec le SCMI.

Les résultats de la figure 18 font apparaître un bon comportement du système de commande dans le domaine de vitesses 0-10 m/s. L'erreur statique est nulle, et la réponse est à la fois rapide et amortie. Au delà, la dynamique est assez considérablement ralentie. Ceci s'explique par une moins bonne modélisation à haute vitesse. Nous avons en effet observé, au chapitre 7, que la dynamique du modèle y est plus rapide que celle du processus. Cette observation se trouve confirmée par les résultats de la figure 18 : la vitesse du modèle s'écarte beaucoup plus de celle du processus à haute vitesse, et sa réponse est plus rapide que celle de ce dernier.

Améliorations à apporter.

Le meilleur moyen d'atténuer ces défauts est bien sûr de réaliser une identification plus précise du véhicule à haute vitesse. Cependant, une amélioration du système de commande actuel pourrait consister à définir un modèle de ralliement pour chaque rapport de vitesse. Le modèle de ralliement actuel a en effet été réglé de manière à conférer au système de commande un comportement satisfaisant en moyenne. Une telle modification permettrait de compenser les écarts de modélisation dans chaque zone du domaine de fonctionnement.

III. ASSERVISSEMENT COMPLET EN TOUT-TERRAIN.

Nous montrons à présent des résultats expérimentaux obtenus en asservissement complet, latéral et longitudinal, sans conducteur à bord (mode pseudo-autonome). La trajectoire de consigne est représentée sur la figure 19. Le terrain est inégal (boue et herbe), et présente deux pentes de 30% (indiquées par des chevrons sur la figure 19). La vitesse de consigne a une valeur constante de 1,67 m/s (6 km/h).

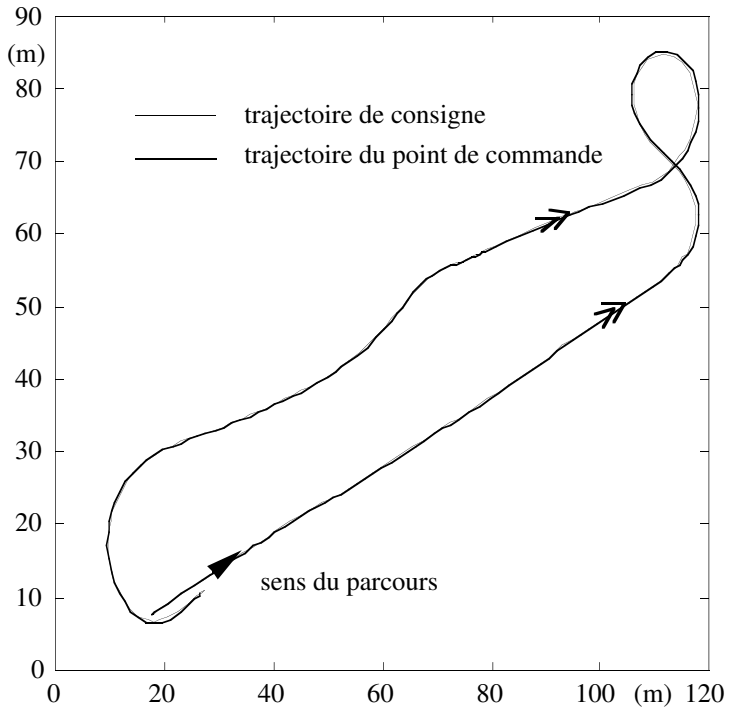

Figure 19.
Résultats expérimentaux d'asservissement complet : trajectoires de consigne et du point de commande.

Les figures 20 et 21 suivantes présentent les résultats des deux asservissements. Toutes les grandeurs sont données en fonction de l'abscisse curviligne le long de la trajectoire de consigne.

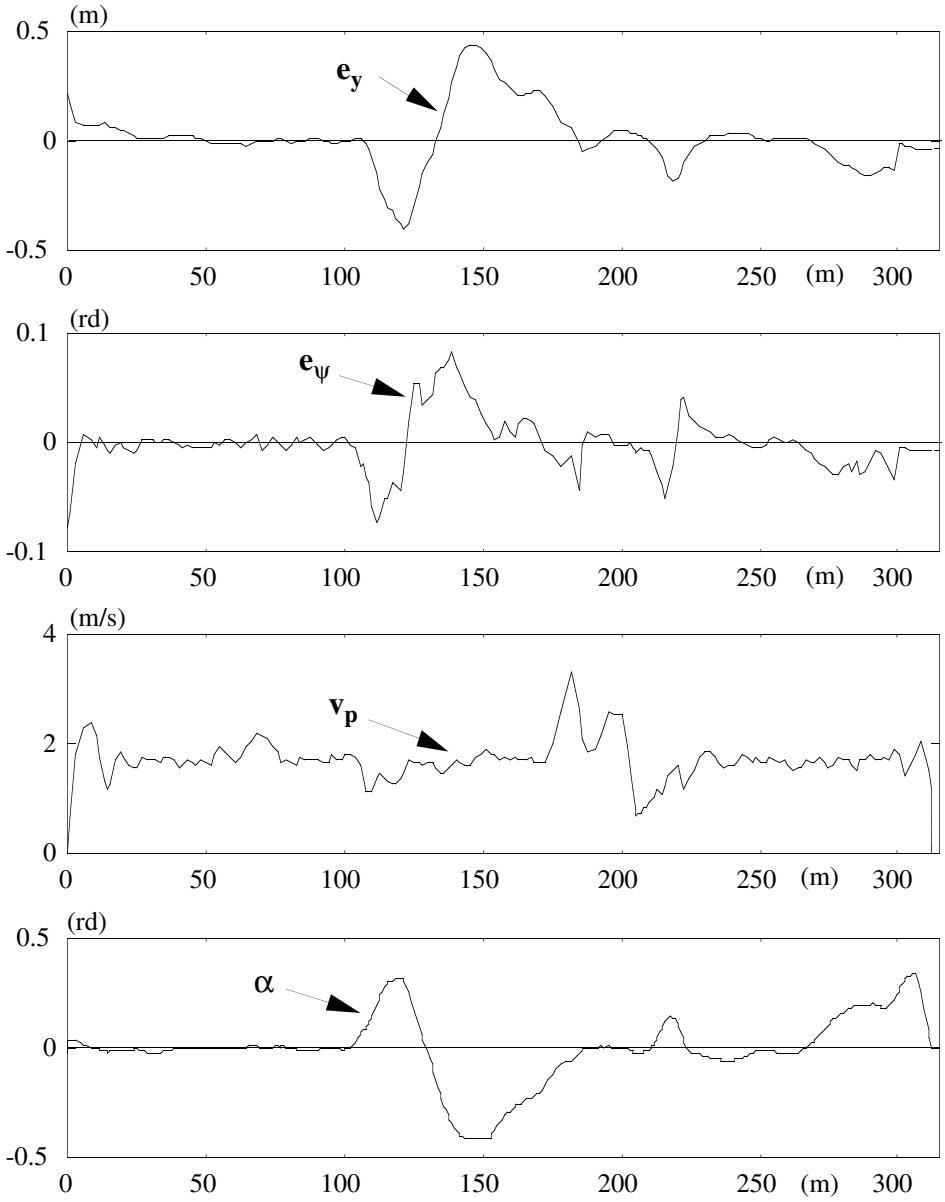

Figure 20.
Résultats expérimentaux d'asservissement complet : asservissement latéral.

Les résultats de la figure 20 font apparaître le maintien des performances du système de commande latéral en tout-terrain : l'erreur latérale ne dépasse pas non plus 40 cm, et ce pour un angle de braquage presque maximal des roues (0.4 rd).

La figure 21 présente les résultats de l'asservissement longitudinal. Les discontinuités de la pente, indiquée sur la courbe du bas, sont bien compensées en montée (instants 50 et 100 s). Pendant la montée (instants 50 à 100 s), l'erreur statique est annulée, pour une valeur de 30% de la pente. Les résultats sont moins bons à la descente (instants 175 à 210 s). Ce défaut pourrait être atténué par une meilleure modélisation du véhicule aux basses vitesses d'une part, et d'autre part par une meilleure prise en compte de la pente par le modèle, qui surestime un peu l'accélération qui lui est due (comme on peut le voir sur l'évolution de la vitesse v du modèle interne).

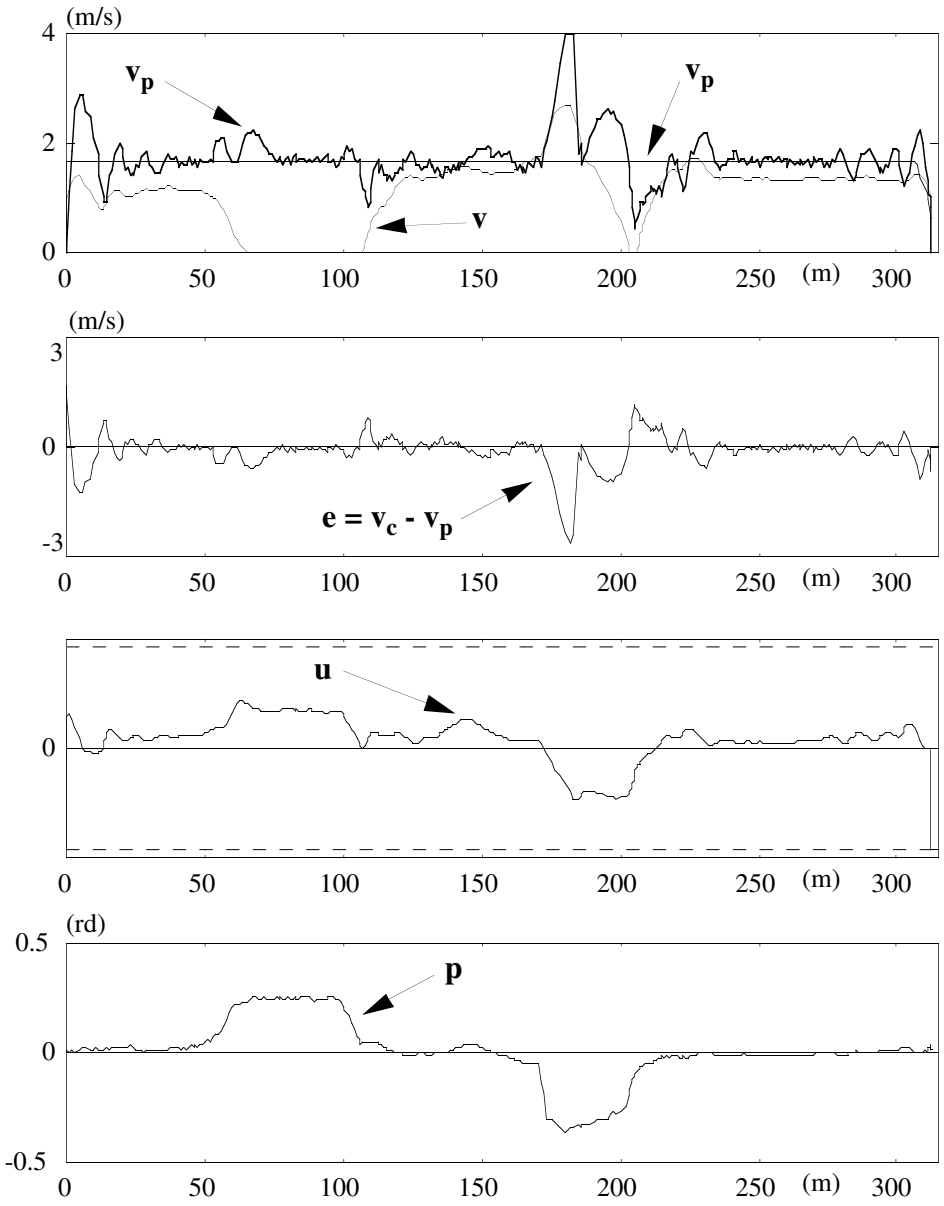

Figure 21.
Résultats expérimentaux d'asservissement complet : asservissement longitudinal.

Notons que le deuxième pic d'erreur de vitesse après le début de la descente (vers 190 s) est dû aux inégalités du terrain à cet endroit (ornières profondes), que révèle d'ailleurs la vitesse du modèle interne, laquelle diminue brutalement au lieu d'augmenter.

Résolution pratique du problème terminal.

Nous avons laissé en suspens le problème de l'arrêt en fin de trajectoire. En effet, aucun système de commande utilisant un correcteur latéral par retour d'état statique invariant ne peut garantir une erreur latérale finale nulle. Dans la pratique, la trajectoire de consigne est prolongée au delà de son point final de manière à pouvoir continuer à définir le point-cible en avant du véhicule selon la règle du §I.2. Lorsque la vitesse de consigne devient nulle (le profil de consigne de la mission en asservissement complet - en mode pseudo-autonome - est conçu en fonction de la décélération maximale du véhicule, d'environ 7 m/s^2), le véhicule s'arrête en pratique très vite, et les erreurs

latérales finales ne dépassent pas la dizaine de centimètres, comme on l'observe sur les figures 8 et 20. Seules des valeurs beaucoup plus importantes de l'erreur justifieraient l'utilisation d'un correcteur d'asservissement sur posture par retour d'état discontinu ou instationnaire à l'approche de la destination finale (correcteurs proposés par [CAN91] [SAM92] [POM92]) comme nous le suggérions dans l'introduction ; de plus ces correcteurs demanderaient à être réglés pour le véhicule.

CONCLUSION.

Le régulateur latéral est un bon exemple de la mise en œuvre de la recherche *ab initio* d'une loi de commande optimale à l'aide d'un réseau de neurones, méthode décrite dans la première partie de ce travail. Les résultats obtenus, aussi bien sur route à fortes variations de courbure (courbure maximale de l'ordre de 0,1 m^{-1}), que sur des terrains accidentés, en dévers (jusqu'à 30%), en pente (jusqu`à 40%), et en tout-terrain avec une pente de 30%, montrent les capacités du système de commande à maintenir d'excellentes performances (erreur latérale inférieure à 40 cm) dans le domaine de fonctionnement non linéaire du véhicule (braquage et vitesse angulaire du volant importants). Les conditions expérimentales extrêmes dans lesquelles ces résultats ont été obtenus rendent difficile la comparaison avec d'autres véhicules automatisés. Ces derniers sont en général testés sur route plane et à faible courbure (0,01 à 0,05 m^{-1}), mais généralement à plus haute vitesse (50-110 km/h), comme par exemple les véhicules décrits dans [JUR93] et [PEN92]. Notre cahier des charges comporte des vitesses de consigne plus faibles (au maximum 72 km/h), mais a le mérite de mettre l'accent sur le comportement non linéaire du véhicule, et de plus en tout-terrain.

De même, l'asservissement longitudinal avec modèle interne et correcteur-S montre la faisabilité de cette approche à l'aide de réseaux de neurones. Malgré les défauts du modèle, le système garantit une erreur statique nulle, même en présence de perturbations importantes (pente de 30%). L'affinement du modèle, et le choix de plusieurs modèles de ralliement pour les différents régimes de vitesse, choix réalisable sans effectuer de nouvel apprentissage, devrait permettre d'améliorer encore ces performances. De plus, l'expérimentation à la SAGEM d'approches classiques apparentées à la commande avec modèle interne et correcteur-S [MIN93b], tend à montrer que ces systèmes de commande sont particulièrement bien adaptés au problème de la conduite en terrain perturbé. Ces considérations nous encouragent à persévérer dans l'investigation et l'amélioration de ces méthodes par l'utilisation de réseaux de neurones.

CONCLUSION GÉNÉRALE

Tout au long du présent travail, nous avons tiré parti de deux caractéristiques fondamentales des réseaux de neurones :

- *la propriété d'approximation universelle parcimonieuse*, dont l'apport est manifeste pour la modélisation et la commande de processus non linéaires ;
- *l'existence d'algorithmes d'apprentissage également universels*, au sens où leur mise en œuvre ne dépend pas de l'application considérée ni de la complexité du réseau soumis à un apprentissage.

Dans un premier temps, nous avons complété le cadre de l'apprentissage des réseaux de neurones élaboré par O. Nerrand pour des réseaux de type entrée-sortie, en l'étendant aux réseaux les plus généraux que sont les réseaux d'état ; nous avons ainsi disposé d'outils algorithmiques complets et puissants pour la modélisation et la commande de processus dynamiques.

C'est en nous appuyant sur ces deux caractéristiques fondamentales des réseaux de neurones que nous avons :

- déterminé les arguments de la fonction théorique que doit réaliser le réseau de neurones selon la tâche considérée (le prédicteur théorique associé à un modèle-hypothèse donné pour une tâche de modélisation, le correcteur théorique associé au modèle d'un processus pour une tâche de commande) : la propriété d'approximation universelle affirme alors que, si cette fonction existe, elle est réalisable par un réseau de neurones ;
- construit un système d'apprentissage du réseau de neurones (éléments constitutifs, algorithme) conduisant à une réalisation de la fonction théorique, si elle existe : l'universalité des algorithmes d'apprentissage garantit que cette réalisation est possible en pratique, quelle que soit la complexité de la fonction.

Dans le domaine de la commande de processus, cette démarche nous a conduit à définir deux familles de systèmes de commande fondés soit sur le correcteur-S, qui impose à un système une sortie de référence, soit sur le correcteur-D, qui lui impose une dynamique de référence. Ces deux familles recouvrent une grande partie des systèmes de commande utilisés en Automatique, et notamment en commande " neuronale ". Nous avons étudié les avantages et les inconvénients respectifs de ces systèmes, en particulier du point de vue de leur robustesse vis-à-vis de défauts du modèle d'une part, et de défauts du correcteur dus à son apprentissage d'autre part. C'est ainsi que nous avons établi les modalités de la mise en œuvre de systèmes de commande neuronaux avec modèle interne.

Par leur conception même, ces systèmes de commande neuronaux avec modèle interne présentent une bonne tolérance aux défauts de modélisation et aux perturbations ; comme il s'agit de systèmes non adaptatifs, le problème de la stabilité des systèmes d'apprentissage ne se pose pas. Néanmoins, ces systèmes de commande ne sont réalisables, avec des réseaux de neurones, que pour des processus dont le modèle est à inverse stable, alors qu'avec des méthodes classiques, il est possible de modifier la synthèse des correcteurs pour éviter cette restriction. Un développement souhaitable consisterait à élaborer les modifications à apporter au système d'apprentissage lorsque le modèle du processus est à inverse instable (comme cela a été fait pour les méthodes de commande avec modèle interne classiques). La commande prédictive neuronale constitue un autre axe de recherche pour la commande de processus à inverse instable.

Tous les exemples illustratifs choisis présentent une dynamique nettement non linéaire excluant une approche linéaire classique, et mettent en lumière l'apport des réseaux de neurones pour une approche non linéaire. L'application industrielle, le pilotage d'un véhicule autonome en tout-terrain, souligne également cet apport. Pour le pilotage latéral, auparavant réalisé à la SAGEM avec des techniques linéaires, la prise en considération des non-linéarités par le correcteur neuronal a permis d'améliorer les performances, sans qu'il soit nécessaire de procéder à aucun réglage après l'apprentissage et la mise en œuvre sur le véhicule. Pour le pilotage longitudinal, aucune technique linéaire ne s'est avérée satisfaisante, et le système de commande neuronal avec modèle interne a obtenu des performances comparables à celle du système non linéaire mis au point à la SAGEM. Ces performances sont très encourageantes dans la mesure où le système de commande n'a fait, ici non plus, l'objet d'aucun réglage supplémentaire. Nous avons ainsi pu vérifier les propriétés de la régulation et de la commande avec modèle interne neuronales, et démontrer la pertinence de l'approche neuronale dans son ensemble sur un problème réel complexe.

L'un des aspects originaux de notre démarche est l'utilisation systématique, pour la conception de modèles et de correcteurs " neuronaux ", des connaissances mathématiques du domaine d'application ; nous avons ainsi tiré parti de l'introduction d'éléments ad-hoc dans les réseaux, éléments déterminés par une analyse physique des phénomènes mis en jeu (modélisation), ou du type de performance souhaité (commande). D'un point de vue conceptuel, nous avons décrit et illustré une approche qui est fondamentalement très voisine de celle de l'automaticien classique. Loin de participer de démarches inconciliables, les synthèses classiques et neuronales de modèles ou de systèmes de commande de processus utilisent des concepts et des méthodes communs, et bénéficient mutuellement de leurs contributions originales respectives.