

# Automatique et Systèmes.

Laurent BARATCHART et Pierre BERNHARD

## 1. Introduction

L'Automatique, comme substantif, désigne généralement un ensemble de problèmes qui se posent à propos de l'automatisation des processus industriels, ou d'autres objets technologiques comme les moyens de transport, les armements, les télécommunications etc. Automatisation veut dire ici, permettre au processus de se dérouler sans intervention d'un opérateur humain.

Pris dans ce sens, le domaine est extrêmement vaste. Il recouvre des problèmes typiquement technologiques et même mécaniques, comme la conception des organes moteurs destinés à actionner les commandes, et physico-mécaniques comme le choix des méthodes de mesure des diverses grandeurs physiques. Suivant la technologie mise en oeuvre, il exige également des connaissances sur les relais, et toute l'électromécanique, l'électronique, tant des courants faibles que de puissance, les automates programmables et les calculateurs de processus. Dans la mesure où nous évoquons les calculateurs numériques, l'Automatique se préoccupe aussi des logiciels temps réel, des systèmes d'exploitation et exécutifs jusqu'aux logiciels d'applications, en passant par les langages spécialisés.

On le voit le champ est vaste. Pourtant ce n'est d'aucun des domaines cités ci-dessus que nous voulons parler ici. Nous allons nous concentrer sur un aspect purement logique, et donc mathématique, de l'automatisation. Un des problèmes de l'automaticien est de *décider* quelle commande appliquer à son procédé, en fonction des informations disponibles: généralement des mesures de grandeurs physiques intervenant dans ce procédé et des connaissances a priori, mais parfois d'un des deux seulement. C'est cette décision que nous voulons approfondir, en tant que processus logique. Ainsi, si la solution d'un problème consiste à doter un arbre moteur et un arbre récepteur d'une poulie chacun, et de relier ces poulies par une courroie, ce que nous retiendrons sera la décision de faire tourner l'arbre récepteur à une vitesse proportionnelle à celle de l'arbre moteur, dans un rapport de proportionnalité donné.

Ce qui nous intéressera sera donc des relations reliant diverses grandeurs logiques, codées, pour les besoins de l'analyse, le plus souvent comme des nombres, mais parfois aussi comme des variables booléennes (c'est à dire prenant leurs valeurs dans l'ensemble  $\{VRAI, FAUX\}$ ), voire des variables plus complexes représentant des connaissances symboliques.

Un *système* sera pour nous un ensemble d'équations reliant ces variables, et que nous utiliserons comme une approximation (une idéalisation) du système physique considéré, ou de la logique de commande que nous souhaitons lui appliquer.

Nous utiliserons le mot de *modèle* dans le sens qu'il a quand on parle de "modèle réduit" pour désigner une maquette. C'est quelque chose qui, d'un certain point de vue, ressemble à l'objet représenté. Un autre point de vue pourrait exiger un autre modèle. L'art de choisir un bon modèle en fonction de la question à laquelle on veut proposer une réponse est un art noble, mais dont nous parlerons peu ici: c'est là que l'*expérience* de l'automaticien joue, et nous savons bien que quand il nous faut nous réfugier derrière ce mot, c'est que nous avons atteint la limite de la connaissance transmissible.

Les questions auxquelles nous nous intéresserons seront la traduction dans le langage de ces équations de questions posées par la pratique de l'automatisation, y compris certains aspects de la modélisation elle-même dans sa partie la mieux quantifiable. (Ainsi que des questions issues du fait que nos modèles ne sont que des approximations.)

Il ressort de cette description que nous considérerons des problèmes de nature mathématique. Il reste cependant une différence fondamentale avec les mathématiques "pures" ou traditionnelles. Science appliquée, l'Automatique n'a résolu un problème que quand elle a donné un moyen de *calculer* une solution. Ainsi les aspects algorithmiques joueront un rôle fondamental, et motiveront eux même les développements mathématiques. Naturellement, cet aspect est tributaire d'une technologie: celle du calcul. Le passage du calcul analogique au calcul numérique a provoqué un renouveau profond de la discipline, et on est encore très loin d'avoir pleinement exploré les possibilités offertes par le calcul réparti ou parallèle.

Nous décrirons ainsi un corpus de connaissances qui s'apparente fortement, par certains aspects, à la recherche opérationnelle et à la théorie de la décision. D'autres aspects appartiennent en propre à l'Automatique et ne lui sont disputés par aucune autre discipline. Nous ne nous intéresserons pas ici aux problèmes de frontière.

Il en est pourtant un où il nous faut prendre un parti, c'est sur la place du traitement du signal. Il s'agit manifestement d'une discipline à part entière, dont les liens avec l'Automatique sont très étroits, d'une part parce que l'automatisation d'un procédé inclut des mesures, donc un signal à traiter, mais surtout par la nature des modèles utilisés dans l'une et l'autre discipline. Nous choisirons ici de ne pas développer le traitement du signal en lui-même, mais seulement ses parties qui ont une forte relation logique avec le reste de notre propos.

Nous avons évoqué la parenté de l'Automatique avec le traitement du signal par la nature des modèles mathématiques utilisés. C'est en fait un phénomène beaucoup plus général. En nous intéressant aux *équations* représentant les processus –et plus précisément à des familles d'équations– c'est à dire à une abstraction, nous traitons du même coup de tous les objets pouvant être représentés par ces équations –ou par des équations de ces familles. Nous avons donc choisi le créneau le moins dépendant de la technologie, le plus stable dans le temps, mais aussi le plus universel. Il est certain que ces techniques sont susceptibles d'être utiles, et ont effectivement été utilisées, dans un domaine très vaste, allant de l'économie et la gestion jusqu'à la dynamique des populations et les sciences de l'environnement, et débordant en tout état de cause largement du cadre étroit de la commande des processus qui les avait vues naître.

On fait parfois remonter la technologie du *feedback*, dont il est largement question ci-dessous, aux clepsydres antiques, qui pourraient avoir ressemblé, dans leur conception, aux robinets à flotteur de nos carburateurs. Outre que cette interprétation des commentaires de Papus soit discutée, ceci ne constituerait de toutes façons pas le départ historique de l'objet de cet article, puisque comme nous l'avons souligné notre propos concerne *l'analyse mathématique* des processus de régulation automatique. Il semble donc qu'il faille faire remonter l'origine de la discipline aux travaux de James Watt, à la fin du 18<sup>me</sup> siècle. Mais il fallut attendre les débuts de l'électronique pour que la discipline se développe vraiment, essentiellement en liaison avec des problèmes de traitement du signal et de conception des amplificateurs à feedback. Les noms de Nyquist, Bode, Shannon, Wiener, sont attachés à ce développement, qui devait conduire à la théorie des asservissements. Pour remarquable que soit cette théorie, le choix que nous avons fait dans cet article est de présenter la théorie moderne des systèmes, pour l'essentiel postérieure à 1960. <sup>(1)</sup>

La fin des années 50 a vu l'arrivée d'un nouvel outil, transformant en profondeur les applications des mathématiques : l'ordinateur. La possibilité de faire des calculs, tant non linéaires que linéaires, devait avoir une influence déterminante sur de nombreuses disciplines, dont la nôtre. A la même époque, la course à l'espace donnait des problèmes, et des moyens, nouveaux aux automaticiens. C'est à cette époque que l'automatique est apparue comme une discipline en propre, distincte de l'électronique, ou plutôt de *l'electrical engineering* suivant la désignation anglosaxonne, qui l'avait abritée jusque là, et devenant une discipline mathématique plus que physique.

Nous renonçons ici à citer les noms des personnes qui ont été les acteurs de ce grand virage. Il faudrait nous résoudre à attribuer les antériorités : ce n'est ni l'objet ni le style de cet article, qui se veut plus didactique que scientifique. Aussi bien, la brève bibliographie que nous donnerons in fine n'a-t-elle pour ambition que d'indiquer au lecteur quelques ouvrages, généralement assez didactiques, dans lesquels compléter son information. Le choix des ouvrages cités, par sa brièveté même, est nécessairement partiel et arbitraire, et reflète plus l'habitude des auteurs qu'un quelconque palmarès.

## 2) Concepts et objectifs de l'Automatique

### 2.1 Systèmes dynamiques

C'est le concept central de l'Automatique. Un *système dynamique* (que nous abrègerons désormais en système) est une application qui, à une *entrée* ou *commande* notée  $u$ , associe une *sortie* notée  $y$ . Reste à définir, bien sûr, ce que sont les entrées et les sorties et c'est ici qu'apparaît le caractère dynamique puisqu'on considérera  $u$  et  $y$  comme des fonctions du temps. Ce temps, à vrai dire, peut revêtir principalement deux formes, à savoir le temps *discret* qui peut être indexé par l'ensemble des entiers, et le temps continu qui peut être indexé par l'ensemble des réels. On parle alors respectivement de systèmes discrets, ou échantillonnés, et de systèmes continus. Dans un calculateur, par exemple, les opérations n'ont lieu qu'à des intervalles

---

<sup>(1)</sup> On en trouvera un exposé dans l'excellent article *Automatique* de Pierre VIDAL dans cette même collection

de temps fixes réglés par l'horloge interne de la machine, de sorte que les modèles afférents seront nécessairement échantillonnés (réciproquement, on peut d'ailleurs affirmer que l'apparition desdits calculateurs est pour beaucoup dans l'essor des modèles discrets). Les lois de la physique, cependant, considèrent le temps comme continu, engendrant des modèles du même type.

Dans la suite, et compte tenu des impératifs de volume, nous aborderons essentiellement les systèmes continus qui sont plus proches, nous semble-t-il de l'ingénierie traditionnelle. Ainsi, tous les systèmes seront désormais supposés continus sauf mention explicite du contraire, mais signalons tout de même que les théories discrètes et continues présentent aussi bien des différences notables que des similitudes étranges. On pourrait penser que la dénomination *temps* est assez arbitraire, et que de nombreux concepts pourraient s'y substituer. Cela est sans-doute vrai, mais le vocable adopté sous-tend l'intuition d'une contrainte majeure et responsable à elle seule d'une grande partie de la complexité de la théorie: le temps est doté d'une dynamique propre *irréversible*.

Notre définition est encore bien vague, puisque nous n'avons pas encore précisé la nature des entrées et des sorties: on conçoit à ce stade que la plupart des concepts de notre univers sont des systèmes. En accord avec le parti que nous avons pris de choisir parmi les aspects les plus concrets et les plus illustratifs de la théorie, nous conviendrons que l'entrée (resp. la sortie) évaluée à l'instant  $t$  donnée par  $u(t)$  (resp.  $y(t)$ ) est un vecteur à  $m$  (resp.  $p$ ) composantes réelles. Cette restriction, ajoutons-le, permet de rendre compte d'un grand nombre de situations pratiques. (Nous évoquerons quelques généralisations nécessaires pour certaines applications).

La difficulté de la théorie des systèmes provient du fait que la valeur de  $y(t_0)$  dépend en général de toutes les valeurs  $u(t)$ . Dans la plupart des modèles, cependant,  $y(t_0)$  ne dépend que des valeurs  $u(t)$  pour  $t \leq t_0$ . Cette propriété est appelée *causalité*, car elle dit *grosso modo* que la sortie est la conséquence de l'entrée. Nous nous limiterons dans ce qui suit aux systèmes causaux, qui recouvrent la majeure partie (mais pas tout à fait la totalité) des situations.

Avec cette dissymétrie naissent en fait les problèmes. Si on commence à s'intéresser à un système  $u \rightarrow y$  à partir de l'instant  $t_0$ , les sorties  $y(t)$  pour  $t > t_0$  n'ont pas du tout le même sens que celles pour lesquelles  $t \leq t_0$ . On peut tenter d'influer sur les premières en choisissant bien  $u(t)$  pour  $t > t_0$ , mais on ne peut que constater les secondes puisque la causalité a pour effet que "ce qui est fait est fait". Et pour influer de manière prédictible sur les sorties futures, encore faut-il pouvoir estimer l'effet qu'auront à l'avenir les valeurs  $u(t)$  pour  $t < t_0$  que l'on n'a pas choisies. D'ailleurs, en poussant plus avant cette manière de voir, on peut même prétendre qu'y compris dans le cas où on aurait choisi l'entrée  $u$  depuis le passé le plus reculé, les imperfections des modèles et les multiples erreurs qui se sont produites à tous les niveaux font que l'on ignore la valeur exacte de l'entrée passée, et que l'on se trouve perpétuellement dans la situation précédente. La description du système en termes de correspondance  $u \rightarrow y$  est peu capable de rendre compte de ce point de vue, qui nécessite l'introduction d'un concept fort important en Automatique à savoir celui d'état. L'état à l'instant  $t_0$  est en quelque sorte l'information minimale qu'il faut posséder pour résumer le passé du système, et connaître l'influence des entrées passées sur les sorties futures. Ceci ne fait évidemment progresser que si l'état peut se calculer à l'aide des entrées futures, et ce plus rapidement qu'il ne change. En outre, son utilisation ne devient calculatoirement effective que si l'information qu'il contient est, en un certain sens, finie. Ceci explique l'importance de ce qu'on appelle parfois les *systèmes à représentation interne de dimension finie* que nous ne chercherons aucunement à décrire en général mais dont nous donnerons le prototype qui est une équation différentielle commandée de la forme

$$\dot{x} = f(x(t), u(t), t), \quad y(t) = h(x(t), u(t), t), \quad (2.1)$$

$$x(t) \in \mathbb{R}^n, \quad u(t) \in \mathbb{R}^m, \quad y(t) \in \mathbb{R}^p.$$

L'état à l'instant  $t_0$  est ici la condition initiale de l'équation, c'est à dire la valeur  $x(t_0)$  qui est en effet une quantité d'information finie ( $n$  nombre réels) qui suffit pour déterminer  $y(t)$  pour  $t > t_0$  si l'on connaît  $u(t)$  pour  $t \geq t_0$ . C'est de classes particulières de tels systèmes que traite en réalité le présent article.

Une dernière notion que nous discuterons est celle de stationnarité. Un système est dit stationnaire si un retard dans l'application de la commande  $u$  produit un retard équivalent de la sortie  $y$ . De manière formelle, cela signifie que si la sortie  $y(t)$  est associée à l'entrée  $u(t)$  et si  $\tau$  est un réel, le système associe à l'entrée  $u(t - \tau)$  la sortie  $y(t - \tau)$ . Cette propriété a des effets simplificateurs manifestes, quoique limités. Le

système (2.1), par exemple, est stationnaire lorsque  $f$  et  $h$  ne dépendent pas de  $t$ . Cependant, les modèles non stationnaires ne sont pas rares, et nous serons amenés à les considérer.

## 2.2 La problématique de l'Automatique et le rôle du feedback

Face à un système, l'automaticien joue en principe un double rôle: il est acteur parce qu'il choisit l'entrée  $u$ , et spectateur parce qu'il observe la sortie  $y$ . Cela lui permet notamment de juger de l'efficacité de l'entrée appliquée, et de la corriger si besoin est. Notre propre corps, par exemple, procède ainsi pour assurer sa station debout.

De fait, concernant un système dynamique donné  $u \rightarrow y$ , discret ou continu, il est essentiellement deux types de problèmes envisagés. Le premier est de choisir  $u$  de sorte que  $y$  possède certaines propriétés, et on dit alors qu'il s'agit d'un problème de *commande*. Le second est d'observer  $y$  pour en déduire certaines choses sur  $u$  comme, le plus souvent, calculer l'état dans lequel le système a été mis. On dit qu'il s'agit d'un problème d'*estimation*, qui prend dans un contexte stochastique le nom de *filtrage*. On peut noter que ces questions concernent, d'une certaine manière, l'inversion de l'opérateur  $u \rightarrow y$ , la première à droite et la seconde à gauche. Ces deux points de vue se conjuguent pour poser le problème de la *synthèse* ou *régulation*: déterminer la commande  $u(t)$  à appliquer à l'instant  $t$  par une observation de la sortie passée. Cela revient en fait à construire un *autre* système dynamique qui, soumis à l'entrée  $y$ , délivre en sortie la commande qui produira les effets désirés. Cette fermeture de la boucle est appelée *feedback*. La conception de cette machine qui marche seule constitue un problème central en Automatique, suffisamment bien posé pour susciter la recherche, suffisamment appliqué pour motiver les résultats, suffisamment général pour n'être jamais véritablement résolu.

Un exemple extrêmement classique de cette problématique, et qui sera repris par la suite, est la stabilisation. Supposons que l'on souhaite astreindre un système stationnaire du type (2.1) à rester dans une position d'équilibre  $x_0$  telle que  $f(x_0, 0) = 0$ , pour maintenir la sortie constante et donc égale à  $h(x_0, 0)$ . Si des perturbations imprévues amènent l'état à la valeur  $x_1$  à l'instant  $t_1$ , il va falloir appliquer une commande  $u(t)$  pour  $t \geq t_1$  afin de ramener l'état à sa valeur  $x_0$ , du moins asymptotiquement. Le concept de feedback consiste ici à choisir pour commande  $u(t)$  une fonction  $U(x(t))$  qui soit telle que l'équation différentielle autonome

$$\dot{x} = f(x, U(x))$$

possède un équilibre stable en  $x_0$ . Le problème est ensuite d'estimer  $x(t)$  à partir de l'observation de  $y(t)$ . Ceci amène à concevoir un système dynamique  $y \rightarrow \hat{x}$ , appelé observateur puisqu'il "observe" l'état, tel que  $x(t) - \hat{x}(t)$  tende vers zero assez vite lorsque  $t \rightarrow \infty$  pour que l'équation différentielle

$$\dot{\hat{x}} = f(\hat{x}, U(\hat{x}))$$

soit encore stable. La correspondance  $y \rightarrow U(\hat{x})$  est un système dynamique nommé compensateur qui, à partir de l'observation de la sortie du système initial, produit en sortie une fonction du temps que l'on peut appliquer à l'entrée du système initial pour obtenir le comportement désiré. Le même schéma peut s'appliquer au suivi d'une trajectoire de consigne où  $u_0$  et  $y_0$  sont maintenant des fonctions non constantes du temps. Naturellement, ces considérations méthodologiques ne suffisent pas à résoudre le problème de déterminer  $U$  et  $\hat{x}$ , problème en général fort difficile et essentiellement local, qui se complique encore davantage lorsqu'on doit tenir compte de l'existence de perturbations affectant  $u$  et  $y$  que l'on modélise par des processus stochastiques. On verra cependant que pour les systèmes linéaires et stationnaires, il existe des réponses substantielles à ces questions qui comptent certainement parmi les réalisations les plus remarquables de l'Automatique, tant par leur simplicité que leur efficacité.

A ce stade, il importe de remarquer que l'application pratique de tout ce qui précède s'appuie sur l'hypothèse qu'on dispose d'un modèle quantitatif précis du système physique mis en jeu. Cependant, dans bien des cas, on ne *connait pas bien* le système que l'on veut commander, soit que la modélisation soit trop complexe, soit que les paramètres en soient mal connus, où encore le plus souvent les deux. On tente alors de construire, un modèle qui explique convenablement les résultats expérimentaux dont on dispose. Nous évoquons ici un domaine de l'Automatique que nous n'avons pas mentionné jusqu'à présent, connu sous le nom d'*identification*. Comme on l'a souligné plus haut, identifier un système dynamique, c'est à dire le décrire par un modèle quantitatif, est un prérequis pour toute étude ultérieure mais constitue un

problème ardu, même pour les modèles les plus simples, et que nous n'aborderons guère. Combinant tous les types de problèmes, la *commande adaptative* consiste à essayer de construire tout à la fois le modèle et le compensateur à partir de l'observation de la sortie  $y(t)$ . Point culminant, en un sens, de la théorie des systèmes, l'accumulation des difficultés a aussi pour effet de limiter quelque peu pour l'instant sa généralité.

### 3) Les systèmes linéaires stationnaires

#### 3.1) Introduction

Le titre de ce paragraphe devrait se poursuivre par “en dimension finie déterministes”, et, pour l'essentiel, “en temps continu”. C'est qu'en effet ces systèmes jouent un rôle central dans la théorie. Historiquement, ils étaient bien adaptés à l'étude des circuits électriques. Aujourd'hui, la théorie des systèmes est utilisée dans d'innombrables domaines, où on ne peut plus toujours faire un hypothèse de linéarité. La théorie linéaire n'en continue pas moins à avoir une place centrale. Dans une certaine mesure, le modèle linéaire joue le rôle de premier terme dans le développement de Taylor d'un modèle non linéaire : il rend compte de la “tendance” au premier ordre. Et sa richesse est telle qu'il nous renseigne déjà considérablement sur ce que peut contenir une dépendance dynamique causale entre signaux. Il n'est pas exagéré de dire que toute étude de modèles non linéaires s'appuie sur la théorie linéaire, ceci bien que toute cette dernière puisse être considérée comme une exploitation du principe de superposition, qui ne subsiste pas dans les modèles non linéaires.

Quant au choix de privilégier le temps continu, nous avons dû nous y résoudre pour une raison de volume total. Indiquons seulement qu'une partie de l'esthétique de toute cette construction vient de l'étonnant parallèle entre la théorie à temps continu et celle à temps discret. C'est dans ce chapitre que le parallèle eut été le plus complet.

Mais avant de quitter cette introduction, donnons une justification plus formelle de l'intérêt des systèmes linéaires dans le cas stationnaire. Les techniques classiques de commande consistent à stabiliser un système dont la sortie est la variable que l'on veut garder “petite”. Supposons que notre modèle linéaire

$$\dot{x} = Ax + Bu$$

soit obtenu par linéarisation d'un système non linéaire

$$\dot{x} = f(x, u)$$

en définissant

$$A = \frac{\partial f}{\partial x}(0, 0) \quad B = \frac{\partial f}{\partial u}(0, 0),$$

et qu'il soit stabilisé par une loi linéaire  $u = Fx$ . Alors, par un théorème de Poincaré, cette même loi stabilise aussi le système non linéaire d'origine, *pour des états initiaux suffisamment proches de 0*.

Malheureusement, si ce théorème semble justifier l'utilisation du modèle linéaire dans tous les cas, il ne dit rien sur *ce suffisamment* proche, limitant considérablement la portée pratique de ce résultat.

#### 3.2) Représentation

La première question qu'il convient de régler est de savoir de quelle famille d'objets (systèmes) on va parler, quelles différentes façons on a de les décrire, et quelles relations il y a entre ces différentes descriptions.

Arbitrairement, nous prenons pour définition d'un système linéaire stationnaire une relation d'entrée-sortie définie à l'aide d'une variable intermédiaire  $x$  appelée l'*état*, de la forme :

$$\dot{x} = Ax + Bu, \tag{3.1}$$

$$y = Cx + Du. \tag{3.2}$$

En outre, on se limitera souvent au cas  $D = 0$ , pour une raison que la formule (3.5) ci-dessous fait ressortir.

Nous verrons que la classe de modèles ainsi retenue n'est pas tout à fait la plus générale possible. C'est ici que se manifeste la restriction “en dimension finie”. On appelle en effet *dimension* du système la dimension, que nous noterons  $n$ , du vecteur d'état  $x$ . Cette classe de modèles présente en particulier

l'intérêt de permettre une étude détaillée de la façon dont **plusieurs commandes affectent plusieurs sorties**. L'état représente toutes les variables qu'il faut connaître à un instant donné pour pouvoir prédire (ou calculer) le comportement futur du système connaissant les commandes futures. Typiquement, l'ensemble des positions et vitesses de tous les points d'un système mécanique, mais aussi bien toutes les températures et concentrations sur les plateaux d'une colonne à distiller, les charges des condensateurs et les courants dans les selfs d'un système électrique, etc.

Donnons un exemple élémentaire. Un point pesant de masse  $m$  et soumis à une force  $u$  a une position  $y$  qui varie suivant la loi  $\ddot{y} = u/m$ . La forme d'état consiste à poser, par exemple,  $x_1 = y$ ,  $x_2 = \dot{y}$ . On a alors

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{m} \end{pmatrix} u. \quad (*)$$

Les grandeurs  $y$ ,  $u$ ,  $x_1$  et  $x_2$  peuvent être des scalaires si le point se déplace sur une droite, ou des vecteurs de  $\mathbb{R}^3$  s'il se déplace dans l'espace physique à trois dimensions.

Pour que l'équation (3.1) fasse correspondre une fonction  $x(t)$  bien définie à toute fonction de commande  $u(t)$ , disons continue par morceaux, il faut préciser des conditions initiales. D'une manière générale, il est habituel de spécifier un instant initial  $t_0$ , et un état initial  $x(t_0) = x_0$ , et de ne considérer le système que sur des intervalles de temps  $[t_0, t_1]$ . On a alors

$$x(t) = \exp[A(t - t_0)]x_0 + \int_{t_0}^t \exp[A(t - \tau)]Bu(\tau) d\tau, \quad (3.3)$$

et  $y(t)$  s'en déduit immédiatement par (3.2). On a ainsi défini une relation *affine* entre les fonctions  $u(\cdot)$  et  $y(\cdot)$  définies sur  $[t_0, t_1]$ .

Cependant dans ce chapitre, pour avoir un modèle réellement linéaire (et non affine) et stationnaire, on adaptera ce schéma de la façon suivante. Les fonctions de commande admissibles seront définies sur  $(-\infty, \infty)$ , on supposera toujours qu'il existe un instant  $t_0$  tel que  $u(t)$  soit nul pour  $t < t_0$ , et  $x(-\infty) = 0$ . Ainsi, si  $u(\cdot)$  est localement sommable (ou, plus simplement, continu par morceaux),  $x(t)$  et  $y(t)$  sont bien définis pour tout  $t$ , et donnés par les formules

$$x(t) = \int_{-\infty}^t \exp((t - \tau)A)Bu(\tau) d\tau \quad (3.4)$$

et (3.2). Il est habituel de réécrire la formule pour  $y$  en termes de la *réponse impulsionnelle* du système. Notons  $\Upsilon$  l'échelon de Heaviside défini par

$$\Upsilon(t) = \begin{cases} 0 & \text{si } t < 0, \\ 1 & \text{si } t \geq 0. \end{cases}$$

Nous utiliserons aussi l'impulsion de Dirac  $\delta$  (qui est la dérivée, au sens des distributions, de  $\Upsilon$ ). Introduisons alors la *réponse impulsionnelle* du système,  $h$ , définie par

$$h(t) = \Upsilon(t)C \exp(tA)B + D\delta(t). \quad (3.5)$$

Les formules (3.2) et (3.4) s'écrivent maintenant

$$y(t) = \int_{-\infty}^{+\infty} h(t - \tau)u(\tau) d\tau. \quad (3.6)$$

Dans l'exemple (\*) ci-dessus, on voit facilement que  $h(t) = (t/m)\Upsilon(t)$ .

On a donc un produit de convolution  $y = h * u$ . La réponse impulsionnelle  $h$  est une fonction (ou une distribution si  $D$  est non nulle) matricielle. Son élément  $h_{ij}(t)$  représente la réponse à l'instant  $t$  de la sortie  $y_i$  à une impulsion à l'instant zéro sur l'entrée  $u_j$ .

Face à un produit de convolution, il est classique d'introduire la transformée de Fourier des fonctions considérées. Les automaticiens lui préfèrent la transformée de Laplace, qui à une fonction  $f(t)$  définie sur  $\mathbb{R}$  fait correspondre la fonction  $F = \mathcal{L}f$  d'une variable complexe  $s$ , définie par

$$F(s) = \int_{-\infty}^{\infty} f(t) \exp(-st) dt.$$

Cette fonction n'est définie que pour les valeurs de  $s$  telles que l'intégrale converge. Mais comme tous les signaux que nous manipulerons seront nuls pour les  $t$  grands négatifs, la borne inférieure de l'intégrale ne posera pas de problème de convergence. Et donc, pourvu que ces signaux ne croissent pas plus vite qu'une exponentielle avec  $t$ , pour une partie réelle de  $s$  suffisamment grande positive, l'intégrale converge. On dit que  $+\infty$  est toujours dans la bande de convergence. (On voit là un intérêt de cette transformée par rapport à celle de Fourier: les conditions d'existence pour certains  $s$  sont peu contraignantes).

On trouvera dans l'article *Automatique* de cette collection, par exemple, une présentation de propriétés de la transformation de Laplace. A la différence de cet article, nous avons défini ici la transformée *bilatère*, en intégrant à partir de  $-\infty$ . Ceci parce que nous permettons au support du signal de commencer dans les temps négatifs. La formule fondamentale (24 in loc. cit.) demeure

$$\mathcal{L} \frac{df}{dt} = sF(s). \quad (3.7)$$

Dans cette formule il faut prendre la dérivée au sens des distributions, c'est à dire que si la fonction présente une discontinuité simple en  $t_0$ , la dérivée contient une impulsion de Dirac en ce point, d'intensité égale au saut de  $f$ , dont la transformée de Laplace est une exponentielle. (Une constante si  $t_0 = 0$ ). Si on considère que  $\dot{f}$  désigne la dérivée de  $f$  là où elle est dérivable, mais ignorant le  $\delta$ , on obtient ainsi (23 in loc. cit.)

$$\mathcal{L} \dot{f} = sF(s) - [f(t_0^+) - f(t_0^-)] \exp(-st_0).$$

Introduisons alors les transformées de Laplace  $U$  et  $Y$  des fonctions  $u$  et  $y$  respectivement, et celle de la réponse impulsionnelle,  $H$ , appelée *fonction de transfert* du système. La relation (3.6) devient alors

$$Y(s) = H(s)U(s). \quad (3.8)$$

L'intérêt de cette forme est évident : c'est sa grande simplicité. Regardons par exemple ce qui se passe si la sortie  $y$  de notre système est elle même l'entrée d'un autre système, de réponse impulsionnelle  $g$  et de fonction de transfert  $G$ , dont on appelle  $z$  la sortie (on dit alors que les deux systèmes sont connectés en cascade).

figure 1

Si on ne dispose que de la relation (3.6), on obtient

$$z(t) = \int \int g(t - \tau) h(\tau - \sigma) u(\sigma) d\sigma d\tau.$$

Naturellement, trois systèmes cascades donneraient une intégrale triple, etc. Avec les fonctions de transfert, on a simplement

$$Z(s) = G(s)H(s)U(s). \quad (3.9)$$

Le gain est encore plus considérable si on considère des systèmes associés en feedback :

figure 2

Les formules en réponse impulsionnelle donnent

$$y(t) = \int h(t - \tau) \left[ v(\tau) - \int g(\tau - \sigma) y(\sigma) d\sigma \right] d\tau.$$

Cette formule est implicite, et il ne semble pas y avoir de moyen simple de la rendre explicite. En termes de fonction de transfert, elle donne

$$Y(s) = H(s)[V(s) - G(s)Y(s)],$$

qui se résout en

$$Y(s) = [I + H(s)G(s)]^{-1}H(s)V(s) = H(s)[I + G(s)H(s)]^{-1}V(s). \quad (3.10)$$

Cette formule est explicite. Nous devons naturellement utiliser la notation  $[I + HG]^{-1}$  et non  $1/[I + HG]$  puisqu'il s'agit de matrices. Cependant, à partir du moment où la matrice  $[I + HG]$  n'est pas identiquement singulière, son inverse existe pour presque tout  $s$ , les  $s$  pour lesquelles elle n'est pas définie constituant simplement les pôles de la fonction. Nous avons donné deux formes équivalentes de cette expression pour faire remarquer que certaines identités peuvent ne pas sauter aux yeux en matière de calcul matriciel.

Une remarque importante pour la suite est que si  $G$  et  $H$  sont des matrices *rationnelles*, c.à d. dont les éléments sont des fractions rationnelles en  $s$ , les opérations (3.9) et (3.10) ne font pas sortir du domaine des matrices rationnelles. Ainsi, comme on sait inverser simplement la transformée de Laplace pour les fractions rationnelles, (cf *Automatique*), la formule (3.10) permet de trouver la réponse impulsionnelle, et donc de "résoudre" la formule implicite correspondante.

Nous avons donc deux types de représentation pour nos systèmes. Les représentations (on verra qu'il n'y a pas unicité) *internes* (3.1) (3.2), et les représentations *externes* (3.6) ou (3.8). Avant d'étudier plus avant comment passer des unes aux autres, il nous faut examiner leurs mérites respectifs.

Un avantage important de la représentation interne est son caractère *récuratif*. Si on connaît l'état  $x(t)$  à un instant  $t$ , on n'a plus besoin de connaître les commandes passées  $u(\tau)$ ,  $\tau \leq t$  pour déterminer les sorties futures du système connaissant les commandes futures. Si on veut "remettre à jour" l'état sur un petit écart de temps, il y aura peu de données à manipuler, peu de calculs à faire. Par comparaison, la formule en fonction de transfert demande que la commande soit connue sur tout l'intervalle  $(-\infty, +\infty)$  pour pouvoir calculer sa transformée de Laplace, et donc celle de la sortie. Elle n'est donc pas utilisable dans un calcul de simulation. Et la formule en réponse impulsionnelle exige, pour chaque nouvel instant  $t$  où la sortie est désirée, que l'intégrale soit recalculée, manipulant l'ensemble des données définissant  $u$  depuis le premier instant où elle est non nulle. En pratique, ceci peut être tempéré par la remarque que souvent, la réponse impulsionnelle est pratiquement nulle pour des  $t$  assez grands, limitant le volume des données à garder en mémoire et à manipuler. Néanmoins, Il s'agit toujours là d'un travail considérablement plus lourd qu'avec une représentation interne. Retenons donc que seule une représentation interne se prête à faire des simulations ou autres calculs en temps réel.

D'un autre côté, on verra que l'ensemble des systèmes qui admettent une représentation de la forme (3.1)(3.2) est un sous ensemble strict de ceux qui peuvent être mis sous la forme (3.6) ou (3.8). Il se trouve que cette classe est néanmoins assez riche pour être intéressante.

La représentation interne souffre par contre d'un défaut beaucoup plus sérieux : c'est sa non-unicité. Deux systèmes, c.à d. deux relations d'entrée sortie, définies par (3.6) sont identiques si et seulement si les réponses impulsionnelles sont identiques. Il n'en va pas de même avec les représentations internes. Deux systèmes de la forme (3.1)(3.2) avec des matrices différentes peuvent être identiques, avoir la même réponse impulsionnelle. D'une part cela pose le problème de caractériser l'ensemble des quadruplets  $(A, B, C, D)$  qui correspondent au même système. Mais aussi, cela rend illusoire l'espoir de retrouver de manière unique ce quadruplet à partir du comportement observé du système, lequel ne dépend que de sa représentation externe.

Une première non-unicité évidente vient du fait qu'on peut changer de base dans l'espace d'état sans rien changer au système. Soit  $T$  une matrice  $n \times n$  inversible, et posons  $\xi = Tx$ . Alors, (3.1) et (3.2) deviennent

$$\begin{aligned} \dot{\xi} &= TAT^{-1}\xi + TBu, \\ y &= CT^{-1}\xi + Du. \end{aligned}$$

Donc manifestement, les quadruplets  $(A, B, C, D)$  et  $(TAT^{-1}, TB, CT^{-1}, D)$  définissent le même système.

Mais la non-unicité peut être plus grave que cela, impliquant des représentations de dimensions différentes du même système. Un exemple trivial, mais dont on verra qu'il est important, est le suivant. Soit

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

et le système

$$\begin{aligned} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} &= \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} u \\ y &= (C_1 \quad 0) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \end{aligned} \tag{3.11}$$

Dans ce système, la variable  $x_2$  n'influence  $y$  ni directement ni par l'intermédiaire de  $x_1$ . Elle ne sert donc à rien, et ce système a la même réponse impulsionnelle que le système  $(A_{11}, B_1, C_1)$ , dont la variable d'état est seulement  $x_1$ .

Déterminons la fonction de transfert du système (3.1)(3.2). La façon facile de faire n'est pas de chercher à transformer la réponse impulsionnelle (3.5), mais directement les équations dynamiques, en utilisant la linéarité de la transformation de Laplace, et la formule (3.7) :

$$sX(s) = AX(s) + BU(s) \quad \text{soit} \quad X(s) = (sI - A)^{-1}BU(s)$$

et donc

$$Y(s) = [C(sI - A)^{-1}B + D]U(s)$$

soit enfin

$$H(s) = C(sI - A)^{-1}B + D. \tag{3.12}$$

Dans l'exemple (\*) ci-dessus, on peut soit utiliser cette formule soit transformer directement l'équation  $m\ddot{y} = u$  pour obtenir  $H(s) = 1/(ms^2)$ .

En utilisant la formule de Cramer, on voit que (3.12) définit une matrice rationnelle *propre*, c'est à dire dont les éléments ont des numérateurs dont le degré n'excède pas celui du dénominateur, et pour laquelle un dénominateur commun possible est le polynôme caractéristique  $\chi_A(s) = \det(sI - A)$  de  $A$ . (Il se peut que des simplifications se produisent, permettant d'exhiber un dénominateur commun de plus petit degré). Une réponse à peu près complète à la question des rapports entre description externe et interne est donnée par le résultat suivant

**THÉORÈME 3.1.** Tout système de la forme (3.1)(3.2) admet une matrice de transfert rationnelle propre. Réciproquement, toute matrice rationnelle propre peut être représentée comme la matrice de transfert d'un système linéaire de dimension finie de la forme (3.1)(3.2). Cette représentation, appelée *réalisation* de la fonction de transfert, est de dimension minimale si et seulement si elle est *complètement accessible et complètement observable* (cf. paragraphe 3.3), et dans ce cas elle est unique à un changement de base dans l'espace d'état près.

Une variante de la représentation externe peut encore être obtenue en écrivant la matrice de transfert sous la forme

$$H(s) = \frac{1}{\chi_A(s)}B(s)$$

où

$$B(s) = Ds^n + B_1s^{n-1} + B_2s^{n-2} + \dots + B_n$$

est un polynôme matriciel. Multiplions chaque membre de (3.8) par  $\chi_A(s)$ , il vient

$$(s^n + a_1s^{n-1} + \dots + a_n)Y(s) = (Ds^n + B_1s^{n-1} + \dots + B_n)U(s).$$

Si on se souvient que  $s^k Y(s)$  est la transformée de Laplace de la dérivée d'ordre  $k$  de  $y$ ,  $y^{(k)}$ , on obtient en revenant dans le domaine temporel

$$y^{(n)} + a_1y^{(n-1)} + \dots + a_ny = Du^{(n)} + B_1u^{(n-1)} + \dots + B_nu$$

qui est une forme "ARMA vectorielle". Très utilisée dans le cas monovarié, cette forme a été utilisée dans le cas multivarié en commande "à variance minimum", et en commande adaptative (cf infra, chapitre 6).

### 3.3) Stabilité

Considérons l'équation différentielle homogène

$$\dot{x} = Ax, \quad x(0) = x_0.$$

On sait qu'elle admet une solution unique, de la forme

$$x(t) = \sum_{k=1}^r p_k(t) e^{\lambda_k t}, \quad (3.13)$$

où les  $\lambda_k$ ,  $k = 1, \dots, r$  sont les valeurs propres distinctes de  $A$ , et les  $p_k(t)$  sont des polynômes en  $t$  de degré strictement inférieur à l'ordre de multiplicité de la valeur propre correspondante. Donc, si (et seulement si) les parties réelles  $\mathcal{R}e(\lambda_k)$  sont toutes négatives,  $x(t) \rightarrow 0$  quelque soit  $x_0$ .

La formule (3.13) s'écrit aussi  $x(t) = \exp(At)x_0$ . Donc  $\exp At$  tend exponentiellement vers zéro si (et seulement si)  $\mathcal{R}e(\lambda_k) < 0$ . La formule (3.3) nous apprend que sous la même condition, le vecteur d'état  $x(t)$  restera borné pour toute commande  $u(t)$  ne croissant pas plus vite que polynômialement.

On dit que

**PROPOSITION 3.1.** Le système (3.1)(3.2) est asymptotiquement stable si les valeurs propres de  $A$  ont leur partie réelle négative.

Nous avons évité, jusqu'ici, de donner une définition précise de la stabilité. C'est qu'en général plusieurs sont possibles. Donnons en une :

**DÉFINITION** Un système est dit "EBSB stable" (entrée bornée - sortie bornée) si à toute entrée bornée sur  $(-\infty, +\infty)$ , correspond une sortie bornée sur  $(-\infty, +\infty)$ .

Un résultat facile, qui ne dépend pas de l'existence d'une forme d'état, est le suivant

**THÉOREME 3.2.** Un système linéaire (3.6) est EBSB stable si et seulement si la réponse impulsionnelle est sommable

$$\int_{-\infty}^{+\infty} \|h(t)\| dt < \infty.$$

Ici,  $\|h(t)\|$  désigne une norme d'opérateur de  $h(t)$ , par exemple sa plus grande valeur singulière, ou tout simplement, pour ce test,  $\sum_{ij} |h_{ij}(t)|$ .

Si  $h(t)$  est sommable, en la remettant dans la formule de la transformée de Laplace pour calculer  $H(s)$ , on voit que l'intégrale converge pour tout  $s$  tel que  $\mathcal{R}e(s) \geq 0$ , puisqu'alors  $|e^{-st}| \leq 1$  pour  $t \geq 0$ , et donc  $e^{st}h(t)$  est sommable. (rappelons que  $h = 0$  pour  $t < 0$ ). On dit qu' $H$  est *holomorphe* dans la région  $\mathcal{R}e(s) \geq 0$ . La réciproque ne peut être énoncée simplement que pour les systèmes de dimension finie.

**THÉOREME 3.3.** Un système linéaire de dimension finie est EBSB stable si et seulement si tous les pôles de sa fonction de transfert ont leur partie réelle négative.

La formule (3.12) montre que les  $s$  pour lesquels  $H(s)$  est infinie sont des valeurs propres de  $A$ . On a donc le résultat suivant :

**THÉOREME 3.4.** Si le système (3.1)(3.2) est asymptotiquement stable au sens de la proposition 3.1 ( $\mathcal{R}e(\lambda_k) < 0$ ), il est EBSB stable.

La réciproque demande quelques précautions supplémentaires. Les matrices  $C$  et  $B$  dans la formule (3.12) peuvent "cacher" certains pôles de  $(sI - A)^{-1}$ . Ainsi, le système (3.11) peut être EBSB stable avec une matrice  $A_{22}$  instable. En effet, seul  $x_2$  diverge alors, et il n'influence pas  $y$ .

Les techniques de stabilité et de fonction de transfert sont utilisées d'une façon extrêmement puissante et élégante pour concevoir des régulations de systèmes monovariabiles dans ce qu'on appelle la théorie des asservissements. On en verra un exposé dans l'article Automatique de cette collection. Malheureusement, largement fondées sur des représentations graphiques dans le plan de  $H(s)$ , ces méthodes s'étendent mal aux systèmes multivariabiles. On verra pourtant dans cet articles plusieurs utilisations des méthodes fréquentielles à de tels problèmes.

### 3.4) Commande modale

Une première question naturelle à poser est de savoir si, en choisissant la commande  $u(\cdot)$ , on peut envoyer l'état  $x$  où l'on veut. Précisons cette question.

DÉFINITION La paire  $(A, B)$  est *complètement accessible* sur l'intervalle  $[t_0, t_1]$  si, quelque soit  $x_1$ , il existe une commande  $u(\cdot) : [t_0, t_1] \rightarrow \mathbb{R}^m$ , telle que le système (3.1) initialisé en  $x(t_0) = 0$  excité par  $u(\cdot)$  satisfasse  $x(t_1) = x_1$ .

THÉORÈME 3.5. La paire  $(A, B)$  est complètement accessible sur  $[t_0, t_1]$  si et seulement si

$$\int_{t_0}^{t_1} \Phi(t_1, t) B B' \Phi'(t_1, t) dt > 0. \quad (3.14)$$

Ici,  $\Phi$  désigne la matrice de transition associée à  $A$ , soit dans le cas stationnaire ( $A = \text{constante}$ )

$$\Phi(t_1, t) = \exp[(t_1 - t)A].$$

et  $> 0$ , pour une matrice (symétrique) signifie “positive définie”. Mais la définition et le théorème 3.5 demeurent identiques si les matrices  $A$  et  $B$  sont variables avec  $t$ , disons continues par morceaux.

Par contre, dans le cas stationnaire, la situation est particulièrement simple comme le montre théorème suivant :

THÉORÈME 3.6. La paire stationnaire  $(A, B)$  est complètement accessible sur un intervalle arbitraire si et seulement si

$$\text{rang}[B \quad AB \quad A^2B \quad \dots \quad A^{n-1}B] = n. \quad (3.15)$$

Le test (3.15) est appelé test de Kalman. (Il apparaissait, dans un autre contexte, dans des livres de calcul des variations bien antérieurs à la théorie des systèmes). La matrice  $[B \quad AB \quad A^2B \quad \dots \quad A^{n-1}B]$  est obtenue en mettant “côte à côte” les  $m$  colonnes de  $B$ , puis les  $m$  colonnes de  $AB$ , et ainsi de suite.

On a aussi défini un système *complètement commandable* comme un système tel qu'on puisse ramener n'importe quel état initial à l'origine. Pour les systèmes en temps continu, ces deux notions coïncident.

Nous avons vu l'importance des valeurs propres de  $A$ , ou pôles, pour la stabilité. Une propriété essentielle d'un système complètement accessible et qu'on peut en placer arbitrairement les pôles par feedback d'état. Considérons une commande de la forme

$$u = -Fx$$

on a alors

$$\dot{x} = (A - BF)x$$

et le théorème de la dynamique arbitraire affirme la chose suivante :

THÉORÈME 3.7. Soit  $(A, B)$  une paire complètement accessible, et  $b(s) = s^n + b_1s^{n-1} + \dots + b_n$  un polynôme à coefficients  $b_k$  arbitraires. Il existe une matrice  $F$  telle que le polynôme caractéristique de  $A - BF$  soit  $b(s)$ .

Nous reviendrons sur l'utilisation de ce théorème.

Signalons enfin que si une paire  $(A, B)$  n'est pas complètement accessible, il existe toujours un changement de base dans l'espace d'état, de matrice  $T$ , tel que

$$TAT^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \quad TB = \begin{pmatrix} B_1 \\ 0 \end{pmatrix}. \quad (3.16)$$

C'est à dire que la partie  $x_2$  de l'état n'est pas influencée par la commande. En particulier, si  $x_2(t_0) = 0$ ,  $x_2$  reste identiquement nul pour tout  $t$ . (Théoriquement. Pour un système “réel”, les perturbations feraient toujours “bouger” un peu, de sorte que si, par exemple,  $A_{22}$  est instable, il diverge toujours).

Une deuxième question à examiner est de savoir si l'observation de la sortie  $y(\cdot)$  permet de retrouver la valeur de l'état. En vertu de la formule (3.3), qui est un principe de superposition, si la commande  $u(\cdot)$  est non nulle, on connaît sa contribution à  $y$  et on peut toujours la retrancher. On peut donc se contenter de considérer le cas  $u = 0$ . La seule partie inconnue qui subsiste ne peut être que  $x_0$ . On posera donc en définition :

DÉFINITION La paire  $(C, A)$  est *complètement observable* sur  $[t_0, t_1]$  si dans le système à commande nulle, la connaissance de  $y(t), t \in [t_0, t_1]$  permet de retrouver  $x_0$ .

THÉOREME 3.8. La paire  $(C, A)$  est complètement observable sur  $[t_0, t_1]$  si et seulement si

$$\int_{t_0}^{t_1} \Phi'(t, t_0) C' C \Phi(t, t_0) dt > 0$$

( $\Phi$  est comme dans le théorème 3.5). A nouveau, la définition et le théorème sont encore valables si  $C$  et  $A$  varient avec  $t$ . Dans le cas stationnaire, cela se simplifie en

THÉOREME 3.9. La paire stationnaire  $(C, A)$  est complètement observable sur un intervalle  $[t_0, t_1]$  quelconque si et seulement si

$$\text{rang} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} = n. \quad (3.17)$$

La comparaison des critères (3.15) et (3.17) montre le fait curieux suivant, que nous élevons au rang de théorème pour le mettre en valeur.

THÉOREME 3.10. (de dualité). La paire  $(C, A)$  est complètement observable si et seulement si  $(A', C')$  est complètement accessible.

On peut déduire de cette remarque et du théorème (3.7) que si la paire  $(C, A)$  est complètement accessible, un choix convenable de la matrice  $K$  permet de placer arbitrairement les valeurs propres de la matrice  $A - KC$ . Aussi, la forme standard que nous avons montrée pour une paire  $(A, B)$  non complètement accessible nous indique, par transposition, que si une paire  $(C, A)$  n'est pas complètement observable, elle peut, par changement de base, se mettre sous la forme (3.11). Ceci explique l'apparition des propriétés de commandabilité et d'observabilité dans les questions de représentation. (Cf théorème 3.1).

Voyons maintenant l'utilisation de ce qui précède dans la régulation d'un système. Quitte à soustraire la sortie désirée, nous supposons que le problème posé est de ramener "suffisamment vite" le système au voisinage de zéro, malgré des perturbations diverses.

Supposons qu'on puisse mesurer l'état. Une idée naïve serait de mesurer  $x(t_0)$ , d'utiliser la complète commandabilité pour calculer un commande  $u(\cdot)$  définie sur un horizon  $T$  pour ramener le système à zéro à  $t_1 = t_0 + T$ . On appliquerait cette commande. A  $t_1$ , en raison des perturbations, l'état ne serait pas tout à fait nul. On itérerait le processus. C'est là une mauvaise façon de faire. La façon qui se révèle au contraire efficace, plus rapide et plus simple, est d'appliquer un feedback  $u = -Fx$ , en choisissant  $F$ , grâce au théorème de la dynamique arbitraire, pour rendre  $A - BF$  "suffisamment" stable. Bien sûr, l'amplitude et la bande passante des actionneurs limitent les performances accessibles en pratique. Le choix des pôles est donc un art assez difficile, sur lequel nous n'avons pas le loisir de nous étendre ici (ceci d'autant moins qu'on verra au chapitre 5 une façon plus intuitive de faire pour trouver  $F$ ).

Mais en général, on ne dispose pas de  $x$  mais seulement de la sortie  $y$ . il est naturel de vouloir utiliser l'observabilité pour recouvrer  $x$ , et faire un feedback d'état. A nouveau, on pourrait décrire une façon naïve de faire, qui essaierait de retrouver la valeur de  $x$  aux instants  $t_0, t_0 + T, \dots$ . A nouveau pourtant, nous allons préférer une méthode basée sur la stabilisation d'un système.

Nous définissons un nouveau système, dont la variable d'état  $\hat{x}$  nous sera entièrement connue, car nous l'initialiserons "au mieux" en un  $\hat{x}_0$  connu, et elle sera ensuite gouvernée par une équation différentielle où tout nous est connu:

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + K(y(t) - C\hat{x}(t)). \quad (3.18)$$

On a fait remarquer au paragraphe (3.1) que la forme d'état se prêtait à une simulation en "temps réel". Ici donc, admettons qu'au fur et à mesure que nous observons les sorties  $y(t)$ , nous pouvons les utiliser en entrée de ce nouveau système, et calculer  $\hat{x}(t)$  en temps réel. Nous allons montrer que si on choisit bien la matrice  $K$ ,  $\hat{x}$  sera une bonne estimée de  $x$ . Posons en effet  $\tilde{x} = x - \hat{x}$ . Soustrayons (3.18) à (3.1) en utilisant aussi (3.2) dans (3.18). Il vient

$$\dot{\tilde{x}} = (A - KC)\tilde{x}. \quad (3.19)$$

De sorte que, grâce au théorème de la dynamique arbitraire "dual" du théorème 3.7, on peut choisir  $K$  pour que,  $A - KC$  étant stable,  $\tilde{x}$  tende vers zéro malgré les perturbations. Ce qui limite ici l'amplitude des

gains  $K$  que l'on peut utiliser, c'est que  $y$  est toujours entaché d'erreurs, ou d'un "bruit"  $w$ , qui se retrouve, amplifié par  $K$ , dans le deuxième membre de (3.19).

La question finale est de savoir s'il est prudent d'utiliser  $\hat{x}$  issu de (3.18) dans un feedback d'état conçu pour placer convenablement les pôles de  $A - BF$ , en faisant

$$u = -F\hat{x}. \quad (3.20)$$

La réponse positive découle du résultat suivant.

**THÉOREME 3.11.** (théorème de séparation modale) Le système de dimension  $2n$  constitué par (3.1), (3.2), (3.18) et (3.20) a pour valeurs propres l'union des valeurs propres de  $A - BF$  et de celles de  $A - KC$ .

On a ici un moyen très puissant de concevoir une régulation multivariable, qui se révèle parfaitement efficace et robuste dans la pratique. La pratique conseillée est de choisir  $C$  pour que l'observateur (3.18) soit à peu près dix fois plus rapide que le commandeur  $A - BF$ .

Pour terminer, évoquons rapidement le cas discret. pour le système

$$x(t+1) = Ax(t) + Bu(t), \quad y(t) = Cx(t),$$

la condition de stabilité porte sur les valeurs propres de  $A$  :  $|\lambda_k| < 1$ . Les critères d'accessibilité et d'observabilité (3.15) à (3.17) subsistent inchangés, et bien sûr aussi le théorème, purement algébrique 3.7. L'observateur s'écrit

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) + K(y(t) - C\hat{x}(t)) \quad (3.21)$$

On remarquera que c'est un prédicteur.  $\hat{x}(t+1)$  est prédit à partir des quantités connues à l'instant  $t$ . Ainsi, tandis qu'on applique la commande  $u(t)$ , on dispose du pas de temps de  $t$  à  $t+1$  pour calculer la commande  $u(t+1) = -F\hat{x}(t+1)$  qu'on utilisera à partir de l'instant  $t+1$ .

### 3.5) A propos des systèmes linéaires non stationnaires.

Nous insérerons ici une digression très brève concernant les systèmes linéaires *non stationnaires*, c'est à dire ceux pour lesquels la représentation (3.1) et (3.2) est possible mais avec des matrices  $A(t), B(t), C(t)$  et  $D(t)$  dépendant du temps. C'est là une classe très importante en pratique, beaucoup de systèmes étant non stationnaires à cause de variations plus ou moins rapides des conditions dans lesquels ils fonctionnent. Un exemple assez classique est celui d'un avion qui possède plusieurs régimes nominaux de vol et donc plusieurs modèles que l'on essaye de raccorder entre eux de façon continue. Nous verrons au chapitre 6 que le contrôle optimal des systèmes linéaires non-stationnaires, est tout de même plus aisé que celui de systèmes non-linéaires plus généraux. Hormis cela, ces systèmes ne sont malheureusement pas beaucoup plus faciles à manipuler que les autres et ne se prêtent en général pas aux considérations de ce chapitre. Mentionnons simplement qu'il est des résultats affaiblis de réalisation lorsque les éléments des matrices  $A(t), B(t), C(t)$  et  $D(t)$  appartiennent à des anneaux de fonctions particuliers à savoir des anneaux Noethériens comme les anneaux de polynômes ou leurs quotients et que, pour les systèmes périodiques c'est à dire ceux pour lesquels les matrices ci-dessus sont des fonctions périodiques du temps, il existe des critères de commandabilité. Un problème notable est la stabilisation de ces systèmes. La commande modale ne s'applique pas pour la raison que l'équation différentielle

$$\dot{x} = A(t)x$$

n'est pas asymptotiquement stable vers l'origine sous la seule condition que les valeurs propres de  $A(t)$  gardent des parties réelles négatives. Pourtant, sous certaines hypothèses de constance de rang et de croissance modérée pour  $t$  assez grand, ce résultat redevient vrai. Ceci suppose *grosso modo* que les coefficients des matrices  $A(t), B(t), C(t)$  et  $D(t)$  ne sont pas trop oscillatoires et croissent lentement à l'infini, et donne quelque fondement théorique à une technique couramment employée sous le nom de *gain scheduling* qui consiste à déterminer à chaque instant un gain  $K(t)$  de telle sorte que les pôles de  $A(t) - B(t)K(t)$  soient plus ou moins constant et à partie réelle suffisamment négative. Cette méthode est souvent d'une réelle efficacité pratique.

### 3.6) Problèmes structurels, approche géométrique et polynômiale

Avec l'essor des techniques précédemment décrites s'est aussi enrichie la problématique de la théorie linéaire, et les premiers succès ont suscité des efforts plus ambitieux pour résoudre des questions de nature plus

structurelles que la simple compensation. Citons par exemple celles de savoir s'il est possible de concevoir le feedback de telle sorte que les perturbations affectant l'entrée ne soient pas présentes dans la sortie (rejet des perturbations), ou bien que les entrées soient naturellement découplées par rapport à une partition des sorties (découplage). Ou encore, est-t-il possible de connecter le système dont on dispose en cascade avec un autre système de façon que l'ensemble se comporte de manière prescrite à l'avance ( poursuite parfaite de modèle)? L'étude de tels problèmes a vu à l'émergence d'au moins deux manières de penser les systèmes linéaires, celles que l'on appelle respectivement "l'approche géométrique" et "l'approche polynômiale". Toutes les deux, chacune à leur manière, visent à linéariser certaines questions qui ne sont pas linéaires en termes des coefficients des matrices d'une réalisation, et ont mis à l'honneur dans le domaine la géométrie linéaire et l'algèbre d'une variable. Sans doute est-il artificiel aujourd'hui de les présenter comme formellement distinctes, alors que leur interpénétration mutuelle est évidente. C'est pourtant ce que nous ferons en discutant très brièvement des questions sus-mentionnées.

L'approche géométrique ne retient d'un système que trois espaces vectoriels  $\mathcal{U}$ ,  $\mathcal{X}$ ,  $\mathcal{Y}$  qui sont respectivement les espaces d'entrée, d'état, et de sortie, et quatre applications linéaires

$$B : \mathcal{U} \rightarrow \mathcal{X}, \quad A : \mathcal{X} \rightarrow \mathcal{X}, \quad C : \mathcal{X} \rightarrow \mathcal{Y}, \quad D : \mathcal{U} \rightarrow \mathcal{Y}.$$

Si on choisit des bases pour identifier  $\mathcal{U}$  à  $\mathbb{R}^m$ ,  $\mathcal{X}$  à  $\mathbb{R}^n$   $\mathcal{Y}$  à  $\mathbb{R}^p$  et  $A, B, C, D$  à des matrices de taille convenable, le système sous-jacent est représenté par (3.1) et (3.2) (remarquer que les différents choix de base donnent des systèmes équivalents), mais c'est justement ce qu'on ne fait pas dans l'approche géométrique, sauf lorsqu'on en vient à calculer. La technique consiste ici à caractériser les propriétés que devraient vérifier les solutions éventuelles d'un problème en termes d'union, d'intersection, et d'invariance sous l'action d'applications linéaires de certains sous-espaces de  $\mathcal{X}$ ,  $\mathcal{U}$ , ou  $\mathcal{Y}$ . La non-vacuité des conditions obtenues est algorithmiquement testable puisque nous sommes en dimension finie, et permet de décider si le problème posé a une solution pour le système considéré. A titre d'exemple, la condition de commandabilité (3.15) signifie que le plus petit sous-espace  $A$ -invariant de  $\mathcal{X}$  contenant l'image de  $B$ , que l'on nomme sous-espace de contrôlabilité de la paire  $(A, B)$  et que l'on note  $\langle A, \mathcal{I}mB \rangle$ , n'est autre que  $\mathcal{X}$  lui-même. La condition d'observabilité (3.17) exprime quant-à-elle la nullité du plus grand sous-espace  $A$ -invariant de  $\mathcal{X}$  contenu dans le noyau de  $C$  que l'on note  $\mathcal{K}erC$ .

Les concepts les plus importants introduits dans la théorie géométrique sont probablement ceux que nous définissons ci-après.

**DÉFINITION** Un sous espace  $\mathcal{V} \subset \mathcal{X}$  est dit  $(A, B)$ -invariant s'il existe une application linéaire  $F : \mathcal{X} \rightarrow \mathcal{U}$  telle que  $\mathcal{V}$  soit  $A + BF$ -invariant, *i.e.*  $(A + BF)\mathcal{V} \subset \mathcal{V}$ .

Exprimé différemment, un sous-espace  $(A, B)$ -invariant est un sous-espace de l'espace d'état que l'on peut s'arranger pour ne jamais quitter si on commande convenablement, à condition bien sûr que l'état initial s'y trouve. Il se trouve alors qu'une telle commande peut toujours être choisie sous forme d'un feedback d'état linéaire. Cette condition est en fait équivalente à  $A\mathcal{V} \subset \mathcal{V} + \mathcal{I}mB$ .

**DÉFINITION** Un sous-espace  $\mathcal{R} \subset \mathcal{X}$  est dit  $(A, B)$ -contrôlable s'il existe deux applications linéaires  $F : \mathcal{X} \rightarrow \mathcal{U}$  et  $G : \mathcal{U} \rightarrow \mathcal{U}$  telles que  $\mathcal{R} = \langle A + BF, \mathcal{I}mBG \rangle$ .

Là encore, et bien que ce soit moins évident, la définition s'interprète géométriquement car elle signifie que deux vecteurs de  $\mathcal{R}$  peuvent être joints par une trajectoire elle-même contenue dans  $\mathcal{R}$ . La condition est aussi équivalente à  $\mathcal{R} = \langle A + BF, \mathcal{R} \cap \mathcal{I}mB \rangle$ .

Comme application de ces notions, considérons le système

$$\dot{x}(t) = Ax(t) + Bu(t) + Sq(t), \tag{3.22}$$

$$y(t) = Cx(t), \tag{3.23}$$

où  $q(t)$  est une perturbation inconnue. Le problème du rejet des perturbations consiste à déterminer s'il est possible de choisir un feedback linéaire d'état  $u = Fx$  tel que la sortie  $y$  soit en fait indépendante de  $q$ . L'examen de (3.3) fournit le résultat suivant:

**THÉORÈME 3.12.** Le rejet des perturbations est possible dans (3.22) et (3.23) si, et seulement si, le plus grand sous-espace  $(A, B)$ -invariant  $\mathcal{V}$  contenu dans  $\mathcal{K}erC$  contient  $\mathcal{I}mS$ . Toute application linéaire  $F$  telle que  $(A + BF)\mathcal{V} \subset \mathcal{V}$  définit alors un feedback qui résout le problème.

Notons que nous n'avons pas tenu compte d'une exigence essentielle en pratique, à savoir que le système bouclé soit stable. On dispose aisément de cette contrainte supplémentaire en introduisant les espaces  $(A, B)$ -invariants stabilisables, ce que nous ne ferons pas ici. Tournons nous à présent vers le second problème que nous avons choisi pour illustrer l'approche géométrique, et considérons le système

$$\dot{x}(t) = Ax(t) + Bu(t), \quad (3.24)$$

$$y_i(t) = C_i x(t), \quad i \in \{1, \dots, k\}, \quad y_i(t) \in \mathbb{R}^{p_i}. \quad (3.25)$$

Le problème du découplage consiste à déterminer s'il est possible de concevoir une commande du type

$$u(t) = Fx(t) + \sum_{i=1}^k G_i u_i(t)$$

telle que, dans le nouveau système

$$\begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_k \end{pmatrix} \rightarrow \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_k \end{pmatrix},$$

$y_i$  soit indépendant de  $u_j$  lorsque  $i \neq j$ , sans pour autant voir affecté l'ensemble des valeurs que peut prendre la sortie  $(y_1, \dots, y_k)$ . La solution de ce problème requiert en général des modifications supplémentaires du système (extension de l'état par compensation dynamique) que nous n'exposerons pas ici, et nous nous contenterons de traiter le cas où la sortie est *complète*, c'est à dire où l'application  $C : \mathcal{X} \rightarrow \mathcal{Y}$  qui à  $x$  associe  $(C_1 x, \dots, C_k x)$  est injective (*i.e.* la sortie permet instantanément de connaître l'état).

**THÉOREME 3.13.** Sous l'hypothèse  $\bigcap_{j=1}^k \text{Ker} C_j = 0$ , le découplage est possible dans (3.24) et (3.25) si, et seulement si, pour tout  $i \in \{1, \dots, k\}$ , le plus grand sous-espace  $(A, B)$ -contrôlable  $\mathcal{R}_i$  contenu dans  $\bigcap_{j \neq i} \text{Ker} C_j$  vérifie  $\mathcal{R}_i + \text{Ker} C_i = \mathcal{X}$ . Plus précisément, si  $\mathcal{R}_i = \langle A + BF_i, \text{Im} B G_i \rangle$ , il existe  $F : \mathcal{X} \rightarrow \mathcal{X}$  qui coïncide avec  $F_i$  sur  $\mathcal{R}_i$ , et une telle application  $F$ , jointe aux applications  $G_i$ , résout le problème.

Là encore, nous n'avons rien dit de la stabilité du système bouclé, dont la dynamique est gouvernée par  $A + BF$ . Si  $(A, B)$  est contrôlable, par exemple, on peut en outre imposer cette stabilité. Mentionnons aussi que, pour offrir des réponses plus nuancées à des questions pratiques, la théorie géométrique a développé les concepts de sous-espace presque invariants et presque contrôlables, pour lesquels les exigences des définitions précédentes peuvent être remplies, sinon exactement, du moins avec une précision arbitraire. On résout alors les problèmes correspondants avec un certain degré de tolérance par rapport à une solution exacte.

Des méthodes algorithmiques ont été développées pour exploiter cette théorie, fondées sur des manipulations d'algèbre linéaire : recherche de rang de matrices, de vecteurs indépendants, de noyaux. Elles ont été largement renouvelées par les apports de la théorie polynomiale que nous abordons à présent.

Il s'agit ici d'appréhender les systèmes par leur matrice de transfert, et de faire apparaître certaines de leurs propriétés à l'aide de factorisations adéquates de ces matrices. On formule alors la réponse à une question sous forme de condition ou d'algorithme algébrique concernant les polynômes apparaissant dans ces factorisations. La structure sous-jacente importante est ici celle de *module* sur l'anneau des polynômes en une variable, la notion essentielle celle de *divisibilité*, l'algorithme de base la *division Euclidienne*, cependant que la possibilité d'effectuer des calculs algébriques par ordinateur donne un caractère opératoire à une telle théorie.

Si  $\mathbb{R}[s]$  désigne l'ensemble des polynômes réels en la variable  $s$ , et si  $M$  est une matrice polynomiale dans  $\mathbb{R}[s]^{p \times m}$ , une manipulation fondamentale consiste à construire deux matrices polynomiales carrées  $U$  et  $V$ , de tailles  $p \times p$  et  $m \times m$  respectivement, qui sont *unimodulaires* c'est à dire que leur inverse est aussi polynômial autrement dit que leur déterminant est un nombre réel non nul, et telles que la matrice  $D = U M V$  soit diagonale au sens où ses éléments  $D_{i,j}$  sont nuls si  $i \neq j$ , cependant que  $D_{i,i}$  divise  $D_{j,j}$  si  $i \leq j$  avec la convention que 0 divise 0. La matrice  $D$  satisfaisant ces conditions est unique (à condition de normer ses éléments en les prenant unitaires par exemple), et est appelée la *forme de Smith* de  $M$ , cependant que  $D_{i,i}$  est appelé le *ième invariant* de  $M$ . Le calcul de matrices  $U$  et  $V$  convenables (qui, elles, ne sont pas uniques en général) est effectif dès lors que l'on sait calculer le p.g.c.d. de deux polynômes, ce qui se

fait par l’algorithme d’Euclide. Tout ce qui vient d’être dit dépend en réalité de ce que  $\mathbb{R}[s]$  est un anneau *principal*, c’est à dire intègre et tel que ses idéaux sont engendrés par *un seul* élément. La considération d’autres anneaux que  $\mathbb{R}[s]$  n’est pas une vaine généralisation comme le montre le problème de poursuite parfaite de modèle: considérons l’équation

$$H_1 H = H_2, \quad (3.26)$$

où  $H_1$  est la fonction de transfert du système dont on dispose et  $H$  la matrice de transfert inconnue du “précompensateur” qu’il faut disposer en série avec le système pour obtenir la fonction de transfert  $H_2$  du comportement global désiré. En remarquant que l’ensemble des fractions rationnelles propres est un anneau principal (à vrai dire assez particulier puisque c’est un anneau de valuation discrète dont le seul premier est  $s^{-1}$ ), on peut trouver  $U$  et  $V$  unimodulaires (ce qui signifie ici qu’elles sont propres ainsi que leur inverse) telles que  $H'_1 = UH_1V$  soit sous forme de Smith. Posant  $H' = V^{-1}H$  et  $H'_2 = UH_2$ , on se rend compte que l’équation en  $H'$

$$H'_1 H' = H'_2, \quad (3.27)$$

qui est équivalente à (3.26), est très simple à résoudre: le problème a une solution si, et seulement si, la  $i$ ème ligne de  $H'_2$  est divisible par le  $i$ ème invariant de  $H'_1$  cependant que les lignes restantes, s’il y en a, sont nulles. Ceci contient la condition *a priori* évidente que le rang de  $H_2$  ne peut excéder celui de  $H_1$ . Comme ces invariants (convenablement normalisés) sont des puissances de  $s^{-1}$  qui est le transfert d’un intégrateur pur, on constate que la seule obstruction à la poursuite de modèle est, en un certain sens, le cas de figure où le système  $H_2$  est plus “rapide” que le système  $H_1$ . L’ensemble des solutions de (3.27), lorsqu’il y en a, est immédiatement décrit comme

$$H' = \begin{pmatrix} A \\ B \end{pmatrix},$$

où  $A$  n’est autre que la matrice obtenue après simplification des invariants non nuls de  $H_1$  dans les lignes correspondantes de  $H'_2$  cependant que  $B$  est rationnelle propre de taille imposée (éventuellement nulle d’ailleurs) mais, hormis cela, arbitraire. En partitionnant  $V = (V_1, V_2)$  de façon adaptée, il en résulte que  $H = VH'$  s’écrit

$$H = H_0 + V_2 B. \quad (3.28)$$

Plus proche de la pratique est le problème de poursuite de modèle avec stabilité, où  $H_1$  et  $H_2$  sont stables et où l’on cherche une solution  $H$  elle aussi stable. Aussi l’approche précédente acquiert-elle plus de valeur encore si l’on remarque que l’anneau des fractions rationnelles stables est encore principal, et que tout ce qui précède peut se transposer à ce cas où, il est vrai, le calcul de p.g.c.d. est plus délicat car il requiert la résolution d’équations algébriques. Il est en outre bien naturel de chercher à minimiser la dimension de l’état du précompensateur  $H$  en utilisant la liberté de choix que l’on a sur  $B$  dans (3.28). C’est alors qu’il faut disposer d’un lien plus caché entre matrice de transfert et réalisation, lequel repose, comme on va le voir, sur la divisibilité.

Pour généraliser aux matrices polynômiales la notion familière de divisibilité entre polynômes, on s’y prend comme suit. Si  $M$  et  $N$  sont deux matrices polynômiales de taille  $p \times m$  et  $p \times k$  respectivement, l’ensemble

$$\mathcal{M} = \{Mv + Nw; v \in \mathbb{R}^m[s], w \in \mathbb{R}^k[s]\}$$

est un module sur  $\mathbb{R}[s]$  qui est un sous-module de  $\mathbb{R}^p[s]$ . A l’aide de la forme de Smith, on constate l’existence d’une matrice  $\Lambda \in \mathbb{R}[s]^{p \times r}$  telle que

$$\mathcal{M} = \{\Lambda v; v \in \mathbb{R}[s]^r\}.$$

Comme les colonnes de  $M$  et de  $N$  appartiennent à  $\mathcal{M}$ , il existe deux matrices polynômiales  $S$  et  $T$  de taille  $r \times m$  et  $r \times k$  respectivement, telles que

$$M = \Lambda S, \quad N = \Lambda T,$$

autrement dit  $\Lambda$  divise  $M$  et  $N$  à gauche. De plus, si  $R$  est une matrice de  $\mathbb{R}[s]^{p \times l}$  qui divise  $M$  et  $N$  à gauche, l’ensemble des vecteurs polynomiaux du type  $Rw$  avec  $w \in \mathbb{R}[s]^l$  est un module qui contient  $\mathcal{M}$ , de sorte que  $R$  divise  $\Lambda$  à gauche, et que  $\Lambda$  est, en ce sens, un *plus grand commun diviseur à gauche* pour  $M$  et  $N$ . Il est défini à la multiplication à droite près par une matrice *unimodulaire*. Lorsque  $\Lambda$  est elle-même

unimodulaire, on dit que  $M$  et  $N$  sont *premières entre elles à gauche*. En transposant chacune des étapes précédentes, on définit de même le *plus grand commun diviseur à droite*.

Si, à présent,  $D$  est une matrice polynômiale  $k \times m$ , on peut introduire sur  $\mathbb{R}^k[s]$  une relation d'équivalence en disant que deux vecteurs polynômiaux sont équivalents lorsqu'ils diffèrent d'un élément de  $D\mathbb{R}^m[s]$ . L'ensemble des classes d'équivalence ainsi obtenu hérite naturellement d'une structure de module sur  $\mathbb{R}[s]$ , qu'on nomme module quotient de  $\mathbb{R}^k[s]$  par  $D$ , et qu'on note  $\mathbb{R}^k[s]/D\mathbb{R}^m[s]$ . Si, de plus,  $D$  est régulière, c'est à dire carrée et de déterminant non nul, la forme de Smith montre que ce module quotient est un espace vectoriel de dimension finie sur  $\mathbb{R}$  égale au degré du déterminant. Notons, par ailleurs, que toute matrice rationnelle  $H$  peut se décomposer, et d'une infinité de façons, en un produit  $ND^{-1}M$ , où  $N$ ,  $D$ , et  $M$  sont polynômiales,  $D$  étant régulière de sorte que son inverse est bien définie en tant que matrice rationnelle.

La plupart des applications de l'algèbre polynômiale à la théorie des systèmes reposent de façon ultime sur le résultat suivant.

**THÉORÈME 3.14.** Soit  $H = ND^{-1}M$  une factorisation de la matrice  $p \times m$  rationnelle propre  $H$ , où  $D$  est de taille  $k \times k$ . Définissons quatre applications linéaires

$$\begin{aligned} A : \mathbb{R}^k[s]/D\mathbb{R}^k[s] &\rightarrow \mathbb{R}^k[s]/D\mathbb{R}^k[s] & B : \mathbb{R}^m &\rightarrow \mathbb{R}^k[s]/D\mathbb{R}^k[s] \\ C : \mathbb{R}^k[s]/D\mathbb{R}^k[s] &\rightarrow \mathbb{R}^p & D : \mathbb{R}^m &\rightarrow \mathbb{R}^p \end{aligned}$$

par les formules

$$A(cl(x)) = s.cl(x), \quad B(u) = cl(Mu), \quad C(cl(x)) = [ND^{-1}x]_1, \quad D(u) = H(\infty)u,$$

où  $cl(x)$  est la classe dans  $\mathbb{R}^k[s]/D\mathbb{R}^k[s]$  du vecteur polynômial  $x$  et  $[ND^{-1}x]_1$  le coefficient de  $s^{-1}$  dans la division suivant les puissances décroissantes du vecteur rationnel  $ND^{-1}x$  (lequel coefficient ne dépend effectivement que de  $cl(x)$ ). Pour tout choix de bases dans  $\mathbb{R}^m$ ,  $\mathbb{R}^k[s]/D\mathbb{R}^k[s]$ , et  $\mathbb{R}^p$  respectivement, les matrices associées aux applications  $A$ ,  $B$ , et  $C$  forment une réalisation de  $H$ . Cette réalisation est commandable si, et seulement si, les matrices  $D$  et  $M$  sont premières entre-elles à gauche. Elle est observable si, et seulement si, les matrices  $D$  et  $N$  sont premières entre-elles à droite.

Cette réalisation est dite *réalisation de Fuhrmann*. Un corollaire en est que la dimension minimale de l'état est le degré du déterminant de  $D$  dans toute factorisation irréductible du type précédent. Cet entier est appelé *degré de Mac-Millan* de  $H$ . Pour le mettre en évidence, on impose parfois à  $D$  d'être *réduite en colonne*, c'est à dire que les vecteurs de  $\mathbb{R}^k$  qui sont les coefficients dominants des diverses colonnes sont indépendants. Ceci a pour effet que la somme des degrés des colonnes est égale au degré du déterminant, et peut toujours être obtenu en multipliant  $D$  à droite par une matrice unimodulaire convenable. On peut naturellement définir par transposition la *réduction en ligne*. Pour illustrer ces concepts, revenons sur le problème de la poursuite parfaite de modèle, et multiplions l'équation (3.26) par un polynôme convenable pour obtenir une équation équivalente du type  $P_1H = P_2$ , où  $P_1$  et  $P_2$  sont des matrices polynômiales. En posant  $H = ND^{-1}$ , on est amené à chercher les solutions de l'équation polynômiale

$$(P_1, -P_2) \begin{pmatrix} N \\ D \end{pmatrix} = 0.$$

En cherchant  $D$  sous forme colonne réduite, cela permet de contrôler à la fois la propriété de  $H$  en bornant le degré des colonnes de  $N$ , et le degré du déterminant. Un algorithme effectif de minimisation de ce dernier peut alors être donné en termes de division Euclidienne pour résoudre la question de poursuite parfaite avec degré minimum du précompensateur. Le cas de la poursuite stable est plus délicat.

Notons, pour finir, qu'on a pu montrer que ce problème est équivalent à celui du rejet des perturbations, dont nous avons vu ci-dessus qu'il se résoud bien par la théorie géométrique. C'est cependant sous forme polynômiale qu'il fut analysé en premier, et, réciproquement, les problèmes de rejet des perturbations, de placement de pôles, etc... ont une solution élégante dans la théorie polynômiale.

### 3.7) Stabilité et robustesse; approche $H_\infty$ .

Si l'approche polynômiale exploite l'existence de la fonction de transfert d'un point de vue algébrique, "l'approche  $H_\infty$ " en distille les aspects analytiques, et en particulier leur rapport à la stabilité. La théorie est

née de la volonté d'analyser la robustesse, vis-à-vis d'erreurs de modélisation ou de perturbations exogènes, des correcteurs tels que les méthodes décrites jusqu'ici permettent de les concevoir. Les débuts de la théorie des systèmes ont vu se développer, dans le cas de systèmes à une entrée et une sortie, une utilisation intensive de la géométrie de l'image, par la fonction de transfert du système mis en série avec le compensateur, de l'axe imaginaire du plan complexe (appelée *lieu de Nyquist*), afin de définir des quantités telles que *marge de gain* et *marge de phase* réputées décrire la robustesse du bouclage effectué. Sans considérer ici ces aspects ni les généralisations du critère de Nyquist proprement dit, il est important de souligner que l'approche  $H_\infty$  renoue, dans le cas de multiples entrées et sorties, avec de telles motivations. Le premier point à clarifier est peut-être la terminologie:  $H_\infty$ , ainsi nommé en référence à G.H. Hardy, est l'ensemble des fonctions analytiques dans le demi-plan droit du plan complexe  $\Pi_+ = \{s \in \mathbf{C}; \operatorname{Re}(s) > 0\}$ , et dont le module est uniformément borné sur ce demi-plan. On définit naturellement la *norme* d'un élément  $f \in H_\infty$  par

$$\|f\|_\infty = \sup_{s \in \Pi_+} |f(s)|.$$

En réalité, les seules fonctions analytiques qui nous préoccupent ici sont celles que l'on dit *réelles*, ceci signifiant qu'elles satisfont à  $\bar{f}(s) = f(\bar{s})$ . La raison de cette restriction est que seules ces fonctions peuvent être des transformées de Laplace de signaux réels, et toute fonction d'une variable complexe sera tacitement supposée réelle dans la suite de ce chapitre. Le compagnon de  $H_\infty$  est l'espace de Hardy  $H_2$  des fonctions  $g$  analytiques dans  $\Pi_+$  mais qui, cette fois, vérifient

$$\sup_{x>0} \int_{-\infty}^{\infty} |f(x+iy)|^2 dy < \infty.$$

La norme  $\|g\|_2$  de  $g \in H_2$  est alors définie comme la racine carrée du *sup* précédent. Notons que les définitions précédentes ne disent rien des valeurs que prend la fonction *sur* l'axe imaginaire lui-même. En fait, il est possible de prouver que si  $f \in H_\infty$  (respectivement  $H_2$ ), la limite (non tangentielle) de  $f$  existe presque partout sur l'axe imaginaire et définit une fonction de  $L_\infty(i\mathbb{R})$  (respectivement  $L_2(i\mathbb{R})$ ) dont la norme n'est autre que  $\|f\|_\infty$  (respectivement  $\|f\|_2$ ). Ceci permet de définir la trace de  $f$  sur l'axe imaginaire et, ayant établi que cette trace définit  $f$  de manière unique, d'identifier  $H_\infty$  (respectivement  $H_2$ ) à un sous-espace *fermé* de  $L_\infty(i\mathbb{R})$  (respectivement  $L_2(i\mathbb{R})$ ). Le complémentaire orthogonal de  $H_2$  dans l'espace de Hilbert  $L_2(i\mathbb{R})$  n'est alors autre que l'espace de Hardy  $\bar{H}_2$  du demi-plan gauche  $\Pi_- = \{s \in \mathbf{C}; \operatorname{Re}(s) < 0\}$ , dont la définition se calque sur la précédente. De manière symétrique, on peut aussi définir  $\bar{H}_\infty$  dont l'usage ici sera moindre.

Le lien entre  $H_\infty$  et  $H_2$  se fonde sur la dualité: si  $f \in H_\infty$ , l'application linéaire  $\Lambda_f : H_2 \rightarrow H_2$  définie par

$$\Lambda_f(g) = fg$$

est continue et sa norme est  $\|f\|_\infty$ , ce qui signifie par définition que c'est le plus petit nombre  $K$  tel que l'on ait pour tout  $g \in H_2$

$$\|\Lambda_f(g)\|_2 \leq K\|g\|_2.$$

C'est cette propriété qui induit la définition des espaces de Hardy à valeurs matricielles: l'espace  $H_2^m$  des vecteurs à  $m$  composantes dont chacune est dans  $H_2$  devient un sous-espace fermé de  $L_2^m$  que l'on muni de la norme

$$\left\| \begin{pmatrix} f_1 \\ f_2 \\ \dots \\ f_m \end{pmatrix} \right\|_2 = (\|f_1\|_2^2 + \|f_2\|_2^2 + \dots + \|f_m\|_2^2)^{\frac{1}{2}},$$

et le complémentaire orthogonal de  $H_2^m$  est à nouveau  $\bar{H}_2^m$ . A présent, si  $F \in H_\infty^{p \times m}$  est une matrice de taille  $p \times m$  dont les éléments appartiennent à  $H_\infty$ , l'opérateur  $\Lambda_F : H_2^m \rightarrow H_2^p$  défini par

$$\Lambda_F(g) = Fg$$

est continu, et sa norme est, *par définition*,  $\|F\|_\infty$ . Lorsque  $F$  n'est plus nécessairement dans  $H_\infty^{p \times m}$  mais simplement dans  $L_\infty^{p \times m}(i\mathbb{R})$ , on peut encore définir  $\Lambda_F$  par la même formule mais il envoie à présent  $H_2^m$  dans  $L_2^p(i\mathbb{R})$ . Si on compose avec la projection orthogonale  $P_-$  de  $L_2^p(i\mathbb{R})$  sur  $\bar{H}_2^p$ , on obtient un opérateur

$$\Gamma_F = P_- \circ \Lambda_F : H_2^m \longrightarrow \bar{H}_2^p$$

qu'on appelle *opérateur de Hankel* associé à  $F$ , et qui jouera un rôle important dans la suite. S'il est vrai que toute fonction de  $L_2(i\mathbb{R})$  se décompose, et ce de façon unique, en somme d'une fonction de  $H_2$  et d'une fonction de  $\bar{H}_2$ , il n'est rien de tel dans  $L_\infty$  où  $H_\infty$  n'est pas complémentable. Quoiqu'il en soit, pour les fonctions *rationnelles* qui sont d'usage en théorie des systèmes, la situation est claire: toute fraction rationnelle propre  $r$  peut s'écrire par décomposition en éléments simples  $r_- + r_0 + r_+$ , où  $r_-$  est strictement propre et a ses pôles dans  $\Pi_+$ ,  $r_0$  est propre et a des pôles imaginaires purs, cependant que  $r_+$  est strictement propres et a ses pôles dans  $\Pi_-$ . Il est alors facile de constater que  $r_-$  appartient simultanément à  $\bar{H}_\infty$  et  $\bar{H}_2$ , que  $r_+$  appartient à  $H_\infty$  et  $H_2$ , et que  $r_0$  ne peut appartenir à  $H_\infty$  sans être constante ni à  $H_2$  sans être nulle. Il s'ensuit qu'une fonction de transfert rationnelle de taille  $p \times m$  est dans  $H_\infty^{p \times m}$  si et seulement si le système correspondant est stable, et qu'elle appartient à  $H_2^{p \times m}$  si et seulement si il est stable et *strictement propre*.

Au vu de la simplicité de la conclusion, il est légitime de s'interroger sur l'utilité de la construction. Les espaces de Hardy, cependant, constituent un cadre adéquat dans lequel poser et résoudre les problèmes d'approximation fonctionnelle qui forment le corps de la théorie linéaire de la robustesse. En outre, les développements qui précèdent s'interprètent de manière convaincante du point de vue de la théorie des systèmes. En effet, un théorème de Paley et Wiener affirme que  $H_2^m$  est constitué précisément des transformées de Laplace des signaux vectoriels de carré sommable qui sont nuls pour les temps négatifs, ce qui entraîne qu'un système du type (3.6) qui associe à une entrée d'énergie finie une sortie d'énergie finie a pour norme d'opérateur  $L_2(0, \infty) \rightarrow L_2(0, \infty)$  la norme  $\|H\|_\infty$  de sa fonction de transfert  $H$ . En d'autres termes, on a

$$\forall u(\cdot) \in L_2(0, \infty), \quad \int_0^\infty \|y(t)\|^2 dt \leq \|H\|_\infty \int_0^\infty \|u(t)\|^2 dt,$$

et  $\|H\|_\infty$  est le plus petit nombre pour lequel l'inégalité a lieu. On comprend à ce stade que la théorie pourrait être exposée au niveau de généralité de (3.6), sans supposer les fonctions de transfert rationnelles c'est à dire les systèmes de dimension d'état finie. La rationalité, cependant, est essentielle pour obtenir des calculs effectifs et nous nous en tiendrons à ce cas.

Reprenons le problème de la poursuite de modèle avec stabilité tel que nous l'avons abordé au chapitre précédent, en désignant dans ce contexte les fonctions de transfert par la lettre  $T$  plutôt que  $H$ . Nous avons alors formulé des conditions pour que (3.26) admette une solution. Que faire à présent si elles ne sont pas remplies? Il est légitime d'essayer de résoudre la question, sinon exactement, du moins de manière approchée et, par exemple, de chercher à minimiser  $\|T_1 T - T_2\|_\infty$  lorsque  $T$  varie dans l'ensemble des matrices rationnelles stables et propres. Nous allons même symétriser quelque peu le problème en introduisant une troisième matrice de transfert  $T_3$ , supposée stable et connue, et en cherchant à minimiser  $\|T_1 T T_3 - T_2\|_\infty$ . On cherche dans ce cas à intercaler un système  $T$  entre deux systèmes  $T_1$  et  $T_3$  pour obtenir un comportement global proche d'un système désiré  $T_3$ . Pour fixer les notations, convenons que  $T_1$  et  $T_3$  sont de tailles  $p_1 \times m_1$  et  $p_3 \times m_3$  respectivement. Il s'ensuit que  $T$  doit être cherchée dans l'ensemble  $RH_\infty^{m_1 \times p_3}$  constitué des éléments rationnels de  $H_\infty^{m_1 \times p_3}$ .

Ici s'introduit le premier outil technique. Une matrice  $F \in H_\infty^{p \times m}$  est dite *intérieure* si la relation  $F'(-s)F(s) = I_m$  a lieu presque partout pour  $s$  imaginaire pur (ceci sous-entend  $p \geq m$ ). Dans cette expression,  $F'$  désigne la transposée de  $F$ , et  $I_m$  est la matrice identité d'ordre  $m$ . On notera dans la suite  $\tilde{F}(s)$  la matrice  $F'(-s)$ . Naturellement, si  $F$  est rationnelle, la relation devient vraie au niveau algébrique. La matrice  $G \in H_\infty^{r \times \ell}$  sera dite *extérieure* si l'image de  $GH_2^\ell$  est dense dans  $H_2^r$  (ceci sous entend  $r \leq \ell$ ). Si  $G$  est rationnelle et que  $G(i\omega)$  reste de plein rang pour tout  $\omega$  fini ou infini, cela signifie que  $G$  admet un inverse à droite dans  $RH_\infty^{\ell \times r}$ . Un résultat fondamental est que tout élément de  $H_\infty^{p \times m}$  se factorise en produit d'une matrice intérieure et d'une matrice extérieure, et ce de façon unique à la multiplication près par une matrice orthogonale. Dans le cas d'une matrice rationnelle, les facteurs sont rationnels également et peuvent être calculés par factorisation spectrale. Rappelons, sans détail, que cela signifie mettre une matrice

carrée rationnelle  $M$  telle que  $\tilde{M} = M$  et  $M(i\omega) \geq 0$  pour tout  $\omega$  (au sens des matrices hermitiennes) sous la forme d'un produit  $\tilde{F}F$ , où  $F$  et son inverse sont analytiques dans  $\Pi_+$ , et qu'il existe des algorithmes effectifs pour cela, en particulier lorsque  $M(i\omega) > 0$  auquel cas ceci peut se faire à partir d'une réalisation de  $M$  en résolvant des équations de Riccati dont il est question plus loin dans le présent article. Dans ces conditions,  $F$  est appelé le facteur spectral de  $M$ .

Revenant à notre problème de poursuite, l'idée de base est à présent d'introduire les factorisations intérieure-extérieure de  $T_1$  et de la transposée  $T'_3$  de  $T_3$ , soit

$$T_1 = U_1 O_1, \quad T'_3 = U_3 O_3.$$

Afin de simplifier la discussion, supposons pour commencer  $U_1$  et  $U_3$  carrées ainsi que  $O_1$  et  $O_3$  de rang plein sur l'axe imaginaire. On peut écrire

$$\inf_{T \in RH_\infty^{m_1 \times p_3}} \|T_1 T T_3 - T_2\|_\infty = \inf_{T \in RH_\infty^{m_1 \times p_3}} \|O_1 T O'_3 - \tilde{U}_1 T_2 \tilde{U}'_3\|_\infty, \quad (3.29)$$

ceci provenant du fait que la multiplication par une matrice carrée  $M \in L_\infty(i\mathbb{R})^r$  (ici  $M = \tilde{U}_1$  ou  $\tilde{U}'_3$ ) telle que  $\tilde{M}M = I_r$  est une isométrie de  $L_2(i\mathbb{R})^r$ . Observons que  $O_1 T O'_3$  parcourt  $RH_\infty^{p_1 \times m_3}$  lorsque  $T$  parcourt  $RH_\infty^{m_1 \times p_3}$ , puisque  $O_1$  et  $O'_3$  sont inversibles à droite et à gauche respectivement dans l'espace de Hardy. Posant  $F = \tilde{U}_1 T_2 \tilde{U}'_3$ , nous sommes donc amenés à chercher

$$\inf_{H \in RH_\infty^{p_1 \times m_3}} \|H - F\|_\infty,$$

où  $F$ , qui est dans  $L_\infty^{p_1 \times m_3}$ , n'appartient pas en général à  $H_\infty^{p_1 \times m_3}$  ce qui rend le problème non trivial (le cas  $F \in H_\infty^{p_1 \times m_3}$  est celui où la poursuite parfaite, étudiée au précédent chapitre, est possible). Un tel problème s'appelle *problème de Nehari*, et la possibilité de le résoudre algorithmiquement est une des pierres angulaires de la théorie. On peut prouver qu'une solution existe, pour laquelle la fonction d'erreur est "passe-tout", et on sait le ramener à une question d'interpolation.

Lorsque  $U_1$  et  $U_3$  ne sont pas carrées, on peut toujours les compléter par des matrices  $V_1$  et  $V_3$  de sorte que  $W_1 = (U_1, V_1)$  et  $W_3 = (U_3, V_3)$  soient carrées et intérieures. On peut alors écrire

$$T_1 T T_3 = W_1 \begin{pmatrix} O_1 T O'_3 & 0 \\ 0 & 0 \end{pmatrix} W'_3,$$

de sorte que, posant cette fois  $F = \tilde{W}_1 T_2 \tilde{W}'_3$ , nous sommes amenés à chercher

$$\inf_{H \in RH_\infty^{p_1 \times m_3}} \left\| \begin{pmatrix} H & 0 \\ 0 & 0 \end{pmatrix} - F \right\|_\infty,$$

qui est dit *problème de Nehari à quatre blocs*. Quoique plus difficile, celui-ci est en principe justiciable de techniques similaires. La détermination effective d'un  $H$  optimal, cependant, est délicate.

En pratique, on préfère calculer des solutions presque optimales selon une méthode itérative ne distinguant pas entre facteurs intérieurs carrés ou non et n'exigeant de résoudre qu'une version affaiblie du problème de Nehari. La clé de cette technique, connue sous le nom de  $\gamma$ -itération, est donnée dans le prochain théorème. Introduisons tout d'abord quelques notations. Posons  $Y = (I_{p_1} - U_1 \tilde{U}_1) T_2$  qui est un élément de  $L_\infty^{p_1 \times p_1}(i\mathbb{R})$  et, si  $\gamma$  est un nombre réel plus grand que  $\|Y\|_\infty$ , soit  $Y_\gamma$  le facteur spectral de  $\gamma^2 I_{m_3} - \tilde{Y}Y$ . Soit à présent  $U_\gamma O_\gamma$  la factorisation intérieure-extérieure de  $(T_3 Y_\gamma^{-1})'$ , et posons

$$Z_\gamma = \tilde{U}_1 T_2 Y_\gamma^{-1} (I_{m_3} - \tilde{U}'_\gamma U'_\gamma),$$

qui est de taille, disons,  $r \times m_3$ . Dans le cas où  $\|Z_\gamma\|_\infty < 1$ , soit de plus  $S_\gamma$  le facteur spectral de  $I_r - Z_\gamma \tilde{Z}_\gamma$  et

$$R_\gamma = (S'_\gamma)^{-1} \tilde{U}_1 T_2 Y_\gamma^{-1} \tilde{U}'_\gamma,$$

qui est de taille, disons,  $r \times s$ . Soient enfin  $D_1$  et  $D_\gamma$  des inverses à droite de  $O_1$  et  $O_\gamma$  respectivement.

THÉOREME 4.15. Soit  $\alpha$  la valeur du *min* dans (3.29). On a alors la relation

$$\alpha = \inf\{\gamma; \gamma > \|Y\|_\infty, \|Z_\gamma\|_\infty < 1, \text{ et } \inf_{H \in RH_\infty^{r \times s}} \|R_\gamma - H\|_\infty < 1\}.$$

En outre, si  $\gamma$  est comme ci-dessus et  $H \in RH_\infty^{r \times s}$  est tel que  $\|R_\gamma - H\|_\infty < 1$ , la fonction  $T$  de  $RH_\infty^{m_1 \times p_3}$  définie par

$$T = D_1 S_\gamma H D'_\gamma$$

est telle que

$$\alpha \leq \|T_1 T T_3 - T_2\|_\infty \leq \gamma,$$

c'est-à-dire réalise une valeur du critère arbitrairement proche de l'optimum.

Pour déterminer  $H$  comme dans le théorème ci-dessus, on est donc amené à résoudre une variante du problème de Nehari qui s'énonce ainsi: étant donné  $R$  rationnel dans  $L_\infty^{r \times s}(i\mathbb{R})$  et tel que

$$\inf_{H \in RH_\infty^{r \times s}} \|R - H\|_\infty < 1, \quad (3.30)$$

déterminer effectivement  $H \in RH_\infty^{r \times s}$  tel que  $\|R - H\|_\infty < 1$ . En fait, le théorème suivant permet en principe de décrire *tous* les  $H$  vérifiant cela. Pour l'énoncer, nous aurons besoin d'introduire la notion de *J-factorisation spectrale*: définissons

$$J_{r,s} = \begin{pmatrix} I_r & 0 \\ 0 & -I_s \end{pmatrix}.$$

Un  $J$ -facteur spectral de  $G \in L_\infty^{(r+s) \times (r+s)}$ , s'il existe, sera une matrice  $G_f$  qui appartient à  $H_\infty^{(r+s) \times (r+s)}$  ainsi que son inverse, et telle que

$$\tilde{G} J_{r,s} G = \tilde{G}_f J_{r,s} G_f.$$

On dira alors qu'on a effectué une  $J$ -factorisation spectrale de  $G$ , les valeurs  $r$  et  $s$  étant données par le contexte. Le calcul d'une  $J$ -factorisation s'effectue avec des moyens très semblables à ceux d'une factorisation spectrale ordinaire. Nous aurons seulement besoin de savoir ici que si  $R_1$  est strictement propre et sans pôles dans  $\Pi_-$ , et si en outre il satisfait (3.30), où  $R$  est remplacé par  $R_1$ , la matrice

$$G = \begin{pmatrix} I_r & R_1 \\ 0 & -I_s \end{pmatrix} \quad (3.31)$$

admet une  $J$ -factorisation spectrale.

Maintenant, si  $R \in L_\infty^{r \times s}$  est une matrice rationnelle satisfaisant (3.30), on peut écrire  $R = R_1 + R_2$  où  $R_1$  est un élément strictement propre de  $\bar{R}H_\infty$  et  $R_2$  un élément de  $RH_\infty$ . Il est clair que  $R_1$  vérifie aussi (3.30) où  $R$  est remplacé par  $R_1$ . Si  $G_f$  est un  $J$ -facteur spectral de  $G$  donné par (3.31), qui existe d'après ce qui précède, définissons  $L \in L_\infty^{(r+s) \times (r+s)}$  par la formule  $L = G G_f^{-1}$ . A tout  $F \in RH_\infty^{r \times s}$ , on peut associer deux matrices  $X_1(F)$  et  $X_2(F)$  appartenant respectivement à  $L_\infty^{r \times s}$  et  $L_\infty^{s \times s}$  par la formule

$$\begin{pmatrix} X_1(F) \\ X_2(F) \end{pmatrix} = L \begin{pmatrix} F \\ I_s \end{pmatrix}. \quad (3.32)$$

Lorsque  $\|F\|_\infty < 1$ , on peut montrer que  $X_2(F)$  est inversible dans  $L_\infty^{s \times s}$  (et même, en fait, dans  $RH_\infty^{s \times s}$ ). Il est donc loisible de définir un élément  $S(F) \in L_\infty^{r \times s}$  en posant  $S(F) = X_1(F) X_2(F)^{-1}$ . Avec ces notations, nous pouvons énoncer:

THÉOREME 4.16. Avec les hypothèses et notations précédentes, l'ensemble des  $H \in RH_\infty^{r \times s}$  tels que

$$\|R - H\|_\infty < 1$$

n'est autre que l'ensemble des matrices qui s'écrivent

$$R - S(F),$$

pour un certain  $F \in RH_\infty^{r \times s}$  tel que  $\|F\|_\infty < 1$ .

A l'aide des théorèmes 4.15 et 4.16, et avec les notations qui y sont introduites, on peut résoudre algorithmiquement le problème de poursuite de modèle initial comme suit.

i) Tout d'abord, on calcule  $Y$ . On sait alors que  $\alpha$  est compris entre  $\|Y\|_\infty$  et  $\|T_2\|_\infty$ . On choisit  $\gamma$  dans cet intervalle.

ii) On calcule  $Z_\gamma$ . Si  $\|Z_\gamma\|_\infty \geq 1$ , augmenter  $\gamma$ , tout en le gardant inférieur à la borne supérieure courante pour  $\alpha$  (au premier pas c'est  $\|T_2\|_\infty$ ), jusqu'à obtenir  $\|Z_\gamma\|_\infty < 1$ .

iii) On calcule

$$\mu(\gamma) = \inf_{H \in RH_\infty^{r \times s}} \|R_\gamma - H\|_\infty.$$

On sait d'après le théorème 4.15 que  $\mu(\gamma) < 1$  si et seulement si  $\alpha < \gamma$ . Si c'est le cas, on a une borne supérieure pour  $\alpha$ , qui devient la borne courante. On diminue alors  $\gamma$ , tout en le gardant supérieur à la borne inférieure courante, et on retourne en ii). Sinon, on a une borne inférieure pour  $\alpha$ , qui devient la borne courante. On augmente  $\gamma$  tout en le gardant inférieur à la borne supérieure courante, et on retourne aussi en ii). On s'arrête sur une valeur de  $\gamma$  lorsque la précision de l'encadrement de  $\alpha$  est considérée comme suffisante.

iv) On calcule grâce au théorème 4.16 un  $H$  tel que

$$\|R_\gamma - H\|_\infty < 1.$$

v) On calcule un  $T$  sous-optimal grâce au théorème 4.15.

L'algorithme précédent, à première vue, peut paraître sans intérêt, puisqu'il requiert à l'étape iii) le calcul de  $\mu(\gamma)$ , ce qui ressemble fort à un problème de Nehari que nous voulions justement éviter. En fait, et c'est là la subtilité de ce calcul, il n'est pas trop difficile de calculer *la valeur* du *min*, alors même que son *argument* est malaisé à obtenir. Ceci provient d'un théorème dit de Nehari, précisément, dont l'énoncé, particularisé au cas rationnel, est le suivant.

THÉOREME 4.17. Soit  $R \in L_\infty^{r \times s}(i\mathbb{R})$  une matrice rationnelle. La quantité

$$\inf_{H \in RH_\infty^{r \times s}} \|R - H\|_\infty$$

n'est autre que  $\|\Gamma_R\|$ , la norme de l'opérateur de Hankel associé à  $R$ .

Du fait que  $R$  est rationnelle, l'opérateur  $\Gamma_R$  se trouve être de rang fini, et on peut calculer sa norme à partir des valeurs propres d'une matrice obtenue en résolvant des équations de Lyapounov. Pour rendre effective l'approche ci-dessus, il importe enfin de pouvoir évaluer la norme  $L_\infty$  d'une matrice rationnelle. Ceci ne peut se faire de manière directe, mais s'obtient par un algorithme itératif qui consiste à tester l'apparition de valeurs propres imaginaires dans une matrice réelle dépendant d'un paramètre, laquelle matrice s'obtient à partir d'une réalisation de la matrice rationnelle initiale.

Si, pour finir,  $O_1$  et  $O_3$  ne sont pas de rang plein sur l'axe imaginaire, les choses deviennent plus difficiles. L'existence d'un argument pour le minimum n'est plus garantie, et on doit envisager des résolutions approchées que nous ne décrivons pas ici.

Si nous avons étudié dans un certain détail le problème de poursuite ci-dessus, ce n'est pas seulement à cause de son intérêt propre, mais surtout parcequ'il est équivalent à un problème canonique en théorie de la robustesse dit *problème standard*. Plutôt que de définir abruptement ce dernier, voyons d'abord comment certains problèmes fondamentaux en automatique y conduisent et considérons pour commencer un système défini par sa fonction de transfert  $P$  que l'on boucle avec un compensateur  $K$  comme indiqué sur la figure 3.

figure 3

Le signal  $r$  correspond à une entrée de référence, cependant que  $y$  est la sortie que l'on souhaite réguler. La première fonction du compensateur sera ici d'assurer la *stabilité interne* du système bouclé, ce qui signifie non seulement que la correspondance  $r \rightarrow y$  est stable, mais encore que les sous-correspondances  $r \rightarrow e$ ,  $u \rightarrow y$ , et  $u \rightarrow e$  le sont aussi. La stabilité de ces sous-correspondances, qui s'expriment par un certain nombre de fonctions de transfert déduites des équations du système bouclé mises sous la forme  $(r, u) \rightarrow (y, e)$ , traduit

le fait qu'il n'y a pas de "modes instables cachés" dans le système. Afin que toutes ces fonctions de transfert "cachées" soient propres, il faut en réalité faire d'autres hypothèses comme par exemple que  $P$  est strictement propre, ce que nous supposons dans la suite. Si  $K$  réalise cette stabilité interne, on dira simplement que  $K$  stabilise  $P$ .

Un premier problème important, du point de vue de la robustesse, est de choisir  $K$  de sorte que la stabilité persiste même si le système "réel"  $P_r$  est légèrement différent de  $P$ , que ce soit à cause d'erreurs de modélisation ou bien de modification du comportement avec le temps (usure, etc...). Une manière d'aborder ce problème en termes relatifs, est de supposer que l'on peut écrire

$$P_r = (I_p + \Delta)P,$$

où  $\Delta$  est une fonction de transfert, éventuellement non rationnelle, représentant l'écart relatif entre  $P$  et  $P_r$ , et  $p$  est le nombre de sorties du système. On fait aussi l'hypothèse que  $\Delta$  est assez faible pour le nombre de pôles instables de  $P$  et de  $P_r$ , comptés avec leur multiplicité, soient les mêmes. Enfin, on suppose qu'on connaît *a priori* une borne sur la perturbation qui s'exprime fréquence par fréquence comme  $\|\Delta(i\omega)\| < |v_1(i\omega)|$ , où le symbole  $\|\cdot\|$  désigne la norme usuelle d'opérateur  $\mathbf{C}^p \rightarrow \mathbf{C}^p$ , cependant que  $v_1$  est une fonction rationnelle de  $L_\infty(i\mathbb{R})$ . Notons qu'on fait ici implicitement l'hypothèse que  $\Delta \in L_\infty^{p \times p}(i\mathbb{R})$ . Si  $K$  stabilise  $P$ , il n'est pas trop difficile de se rendre compte (c'est essentiellement un analogue, utilisant les valeurs singulières, du critère de Nyquist) qu'une condition suffisante de stabilité robuste est

$$\|v_1 PK(I_p + PK)^{-1}\|_\infty < 1. \quad (3.33)$$

Un autre aspect important, du point de vue de la robustesse, est que la fonction de transfert du système bouclé soit peu sensible aux variations  $\Delta$  dont nous essayons à l'instant de maîtriser la stabilité. En notant  $T_b = (I_p + PK)^{-1}PK$  la fonction de transfert du système bouclé, la fonction "réelle"  $T_r$  consécutive à une variation relative  $\Delta$  de  $P$  s'écrit

$$T_r = (I_p + \Delta_T)T_b, \quad \text{avec} \quad \Delta_T = (I_p + P_r K)^{-1}\Delta.$$

La fonction de transfert  $\Delta_T$  représente ainsi la variation relative de  $T_r$  sous l'effet des perturbations  $\Delta$ . La dérivée de  $\Delta_T$  par rapport à  $\Delta$  au point  $0 \in L_\infty^{p \times p}(i\mathbb{R})$  vaut  $(I_p + PK)^{-1}$  et s'appelle *matrice de sensibilité* du système. C'est elle que l'on cherche à rendre petite, et en l'espèce majorée en norme d'opérateur, en chaque point de l'axe imaginaire, par le module d'une fonction rationnelle qu'il est plus commode de définir par son inverse  $v_2 \in L_\infty(i\mathbb{R})$ , en demandant aussi que  $v_2^{-1} \in L_\infty(i\mathbb{R})$ . Ceci revient à exiger

$$\|v_2(I_p + PK)^{-1}\|_\infty < 1. \quad (3.34)$$

Naturellement, (3.33) et (3.34) sont, dans une certaine mesure, contradictoires, puisque  $PK(I_p + PK)^{-1}$  et  $(I_p + PK)^{-1}$ , dont la somme vaut  $I_p$ , ne peuvent être simultanément petits (ce qui entraîne en passant que les fonctions  $v_1$  et  $v_2$  ne peuvent être choisies indépendamment). En fait, c'est un des atouts majeurs de la théorie que de permettre de synthétiser un compromis acceptable, en jouant sur le fait que (3.33) et (3.34) ne sont pas effectives dans les mêmes plages de fréquence, car c'est dans les hautes fréquences que se concentrent les erreurs et perturbations  $\Delta$  que (3.33) doit stabiliser, alors que le signal  $r$  est plutôt situé dans les basses fréquences, qui sont donc celles qui concernent (3.34).

Introduisons à présent le diagramme standard de la figure 4

figure 4

où  $G$  est la fonction de transfert du système que l'on doit commander,  $K$  celle du compensateur que l'on doit concevoir,  $w$  une entrée constituée de commandes nominales, perturbations et bruits divers,  $z$  la sortie que l'on cherche à réguler,  $y$  la sortie mesurée, et  $u$  la commande élaborée par le compensateur à partir de  $y$ . Les entrées  $\nu_1$  et  $\nu_2$  représentent, quant à elles, des perturbations sur la commande et l'observation respectivement. Conformément à ce qui précède, on dit que  $K$  stabilise  $G$  si les neuf fonctions de transfert liant  $w$ ,  $\nu_1$ ,  $\nu_2$ , d'une part, à  $z$ ,  $u$ ,  $y$ , d'autre part, sont stables.

Le problème standard associé consiste à déterminer  $K$  stabilisant  $G$  tel que la fonction de transfert  $T_{w,z}$  liant  $w$  à  $z$  soit de norme  $H_\infty$  minimale. Dans notre cas, en posant

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

et

$$G = \begin{pmatrix} v_2 I_p & -v_2 P \\ 0 & v_1 P \\ I_p & -P \end{pmatrix},$$

on constate que

$$T_{w,z} = \begin{pmatrix} v_2(I_p + PK)^{-1} \\ v_1 PK(I_p + PK)^{-1} \end{pmatrix}.$$

Ainsi, l'assertion

$$\|T_{w,z}\|_\infty < 1 \tag{3.35}$$

est logiquement plus forte que (3.33) et (3.34) réunies, et si le choix de  $v_1$  et  $v_2$  autorise (3.35) lorsque  $K$  est solution de ce problème standard, on aura trouvé un correcteur convenable. C'est donc sur la résolution du problème standard que nous nous concentrons maintenant.

Il se peut que ce problème n'admette pas de solution pour la raison que  $G$  n'est pas stabilisable. Nous exclurons ce cas, ce qui revient à supposer que  $P$  est stabilisable. Nous aurons besoin ici d'une factorisation analogue à celle introduite lors de l'approche polynômiale, à ceci près que les matrices polynômiales sont remplacées par des matrices rationnelles stables: toute matrice rationnelle propre peut s'écrire  $NM^{-1}$ , où  $N$  et  $M$  sont rationnelles propres et stables de taille convenable, et de plus premières entre elles à droite au sens où il existe une relation de Bezout:

$$X_1 N + Y_1 M = I,$$

où  $I$  est une identité de taille convenable,  $X_1$  et  $Y_1$  étant encore rationnelles propres et stables. On peut bien sûr écrire aussi une factorisation du type  $M_1^{-1} N_1$ , où  $M_1$  et  $N_1$  sont à présent premières entre elles à gauche. L'existence se démontre pareillement, puisque  $RH_\infty$  est principal. En fait, on peut même choisir  $M$ ,  $N$ ,  $M_1$ ,  $N_1$  de sorte que l'on ait

$$\begin{pmatrix} X_1 & -Y_1 \\ -N_1 & M_1 \end{pmatrix} \begin{pmatrix} M & Y \\ N & X \end{pmatrix} = I, \tag{3.36}$$

où  $I$  est une matrice identité de taille convenable et où  $X$ ,  $Y$ ,  $X_1$  et  $Y_1$  sont des matrices rationnelles propres et stables qui se trouvent être nécessairement, au signe près, des facteurs de Bezout pour  $(N_1, M_1)$  d'une part, et  $(N, M)$  d'autre part. Une telle factorisation est dite *doublement première*, et elle peut se calculer à partir d'une réalisation en effectuant la synthèse spectrale évoquée au premier paragraphe. Son intérêt réside ici dans le fait qu'elle autorise la paramétrisation de tous les compensateurs stabilisant  $G$  dans le diagramme standard, espace sur lequel s'effectue l'optimisation dans le problème standard. En effet si on partitionne  $G$  de façon à pouvoir écrire

$$\begin{pmatrix} z \\ y \end{pmatrix} = \begin{pmatrix} G_{1,1} & G_{1,2} \\ G_{2,1} & G_{2,2} \end{pmatrix} \begin{pmatrix} w \\ u \end{pmatrix},$$

on a le résultat suivant:

**THÉORÈME 4.18.** Si (3.36) est une factorisation doublement première de  $G_{2,2}$ , et si  $G$  est stabilisable, l'ensemble des  $K$  rationnelles propres stabilisant  $G$  dans la figure 4 est paramétré par les fonctions de transfert  $T$  rationnelles propres et stables de taille convenable par les formules:

$$K = (Y - MT)(X - NT)^{-1} = (X_1 - TN_1)^{-1}(Y_1 - TM_1).$$

Si l'on pose en outre

$$T_1 = G_{1,2}M, \quad T_2 = G_{1,1} + G_{1,2}MY_1G_{2,1}, \quad T_3 = M_1G_{2,1},$$

et si  $K$  est défini comme ci-dessus, la fonction de transfert liant  $w$  à  $z$  dans le système bouclé de la figure 3 s'écrit

$$T_{w,z} = T_2 - T_1 T T_3. \quad (3.37)$$

Au vu de ce théorème et, en particulier, de l'équation (3.37), il devient clair que le problème standard se ramène au problème de poursuite de modèle exposé dans la première partie de ce paragraphe. Le problème standard peut supporter maints ajouts et variations de manière à traiter rejet de perturbations, découplage, et autres questions de la théorie que nous avons déjà eu l'occasion d'introduire. Les perturbations peuvent aussi être traitées de façon absolue plutôt que relative quoique, et c'est une limitation de la théorie, elles soient toujours supposées linéaires ce qui signifie qu'on ne peut corriger qu'au premier ordre, ce qui peut influencer notablement sur la marge de stabilité car le domaine de validité de l'approximation linéaire dépend bien entendu de la correction elle-même.

Il a pu apparaître comme une règle, dans ce qui précède, que la théorie analytique est rendue effective par des calculs en variable d'état utilisant de façon essentielle la rationalité. Cela ne signifie pas que les calculs polynômiaux effectués directement sur les fonctions de transfert sont impossibles; ils sont seulement numériquement plus délicats. Il existe d'autres approches du contrôle  $H_\infty$ , qui font jouer un rôle central à la représentation interne et aux équations matricielles et qui évacuent pour l'essentiel les espaces de Hardy. Elles ont permis d'aborder des questions importantes que nous ne traitons pas ici, comme par exemple l'ordre des compensateurs obtenus. L'approche retenue ici, en sus de son importance historique et de l'adéquation de ses concepts, exhibe de beaux liens avec la théorie classique des problèmes extrémaux dans les espaces de Hardy.

#### 4) Commande adaptative

La commande adaptative constitue aujourd'hui un domaine de recherche très vaste, dont les débouchés dans la pratique de l'ingénieur sont plus en promesse qu'actuels, malgré quelques réalisations récentes. Aussi, au bénéfice du volume du présent article, renonçons-nous à en donner une description technique, pour souligner plutôt les problèmes qu'elle aborde et les difficultés qu'elle rencontre. On trouvera au chapitre suivant un exemple, le plus élémentaire possible, de comment elle peut procéder.

La problématique de la commande adaptative est fondamentalement assez peu différente de celle de la commande robuste (non-adaptative). Dans les deux cas, l'objectif est de maintenir un niveau de performances imposé au départ en dépit d'incertitudes sur la dynamique du processus à commander ou sur la façon dont cette dynamique se modifie au cours du temps. C'est au niveau des solutions envisagées que les différences apparaissent: alors que la commande robuste vise à synthétiser un régulateur de structure invariante capable "d'absorber" les incertitudes dans une plage donnée, à partir de l'étude du cas le plus défavorable, la commande adaptative repose sur un principe de mise à jour (estimation récursive) d'un modèle du processus à partir de l'observation de ses entrées et de ses sorties, et de modification (ou "d'adaptation") du calcul de la commande en se basant sur le modèle estimé courant. La formalisation de ce principe et les premières études semblent dater de la fin des années 50, avec des méthodes de moindres carrés récursifs.

Par la suite, les travaux sur la commande adaptative tendent à s'unifier autour des concepts de *Model Reference Adaptive Systems*, initialement développé dans un cadre déterministe pour résoudre le problème de la poursuite de signaux de référence, et de *Self-Tuning Regulators*, (voir chapitre 5), posé au départ dans un cadre stochastique pour résoudre le problème de régulation autour d'une consigne. Les deux concepts sont en fait complémentaires et peuvent être regroupés au sein d'une formulation plus générale de la commande adaptative. (Voir les ouvrages de référence cités in fine)

Les années 80 ont vu les travaux s'orienter vers l'étude théorique des questions centrales de stabilité et de robustesse des schémas adaptatifs. En particulier, la bornitude de l'ensemble des signaux générés par le processus et l'algorithme d'identification/commande a été démontrée dans le cadre d'hypothèses opératoires simplificatrices se résumant, dans leur essence, à: *le système à commander est linéaire invariant et un majorant de l'ordre du système est connu*. L'incertitude motivant l'adaptation provient dans ce cas de la méconnaissance initiale des paramètres de la fonction de transfert du système. Les premières démonstrations étaient limitées aux systèmes à *minimum de phase* et à des stratégies de commande de type *minimum de variance*. Leur extension aux systèmes à *non-minimum de phase* et à d'autres stratégies de commande a été obtenue au prix d'une hypothèse technique (commandabilité ou stabilisabilité du modèle estimé) dont le contournement a donné lieu à des études spécifiques. Les progrès réalisés dans l'établissement de la propriété

de bornitude de l'ensemble des signaux se sont traduits par un amoindrissement progressif des hypothèses opératoires afin de se rapprocher de conditions plus proches de la pratique: perturbations additives bornées, dynamiques rapides non modélisées, dérive des paramètres du système.

La robustesse des performances de la commande adaptative pose cependant un problème de fond lié au fait qu'un système peut, dans certaines conditions, subir de profondes modifications sans que cela soit détectable par le biais de l'observation de ses entrées et de ses sorties, ou sans que cette observation permette de reconstituer entièrement la modification subie par le système. Rien ne permet par exemple de distinguer deux systèmes linéaires stables au repos (entrées et sorties nulles). Il suffit alors que l'objectif de commande soit compatible avec ces conditions (cas de la régulation autour d'une consigne) pour que le mécanisme d'identification ne puisse plus remplir son rôle informatif de façon fiable, ou soit même "trompé" par l'action de perturbations agissant sur le système. Lorsque le calcul de la commande fait "aveuglement" confiance à la qualité du modèle estimé, il peut alors se produire des phénomènes temporaires de forte instabilité inacceptables dans la pratique. L'étude des mécanismes d'apparition de ces phénomènes a fait l'objet de plusieurs travaux.

L'amélioration de la robustesse de la commande adaptative est depuis toujours le principal souci des praticiens du domaine. D'elle dépend sans doute la réussite du transfert de la théorie vers les techniques de l'ingénieur. De nombreuses méthodes, plus ou moins *ad hoc*, ont été proposées dans ce but, mais aucune d'entre elles ne semble jusqu'à présent faire l'unanimité. En particulier, une polémique subsiste quant à la façon de résoudre le problème d'identification évoqué précédemment. Une technique consiste à ajouter aux valeurs de commande des perturbations au contenu fréquentiel riche de sorte à "forcer" le système à fournir en permanence l'information nécessaire à son identification et satisfaire une condition dite *d'excitation persistante*. D'un point de vue pratique, cet ajout peut cependant être indésirable parcequ'incompatible avec les objectifs de commande. La solution est-elle alors un mariage de raison avec la commande robuste?

## 5) La théorie linéaire quadratique

### 5.1) introduction à la théorie linéaire quadratique gaussienne

Il n'est pas exagéré de dire que le succès de la théorie des systèmes en variables d'états, telle que nous l'avons exposée au chapitre 3 par exemple, est en grande partie dû à la théorie linéaire quadratique gaussienne, qui, outre son élégance mathématique, donne un moyen puissant pour régler un compensateur. Dans une certaine mesure, il s'agit d'une version de la méthode des "moindres carrés", encore que, si ceci s'applique clairement au problème *LQ* du paragraphe suivant, quand on explique ainsi le filtre de Kalman, Kalman lui-même, le "père" de toute cette théorie, aime citer Einstein affirmant qu'il faut expliquer les choses de manière aussi simple que possible, mais *pas plus!*

Cette théorie s'est développée au cours des années 1960 à 1970. Les succès spatiaux américains de cette époque, culminant avec les débarquements sur la lune (1969), auxquels ces techniques ont contribué, en ont aussi assuré la publicité qu'elles méritaient assurément. Depuis, leur utilisation s'est largement répandue dans les systèmes technologiques de pointe (aéronautique notamment), puis plus lentement dans d'innombrables autres domaines, allant de la commande des processus chimiques jusqu'à la modélisation économique.

Le paragraphe 5.2 traite du problème de commande, en supposant qu'on dispose d'un modèle "parfait" et qu'on mesure exactement l'état. Le paragraphe 5.3 traite de l'estimation de l'état en présence d'un modèle imparfait (un "bruit de dynamique") et de mesures partielles et imparfaites de l'état, un problème d'estimation intéressant en soi. Le théorème de séparation fonde l'usage "naturel" de ces deux résultats ensemble pour traiter la commande d'un système bruité avec des observations imparfaites.

Nous examinerons ensuite ce que devient cette théorie dans un cadre stationnaire, caractérisant les pôles de l'observateur et du commandeur ainsi construits, pour mieux faire le lien avec le chapitre 3. Quelques extensions compléteront l'image, avant d'aborder avec le paragraphe 5.6 une approche complètement différente du traitement des incertitudes, toujours dans un cadre linéaire-quadratique. Nous commenterons cette nouvelle approche au paragraphe 5.6 lui-même.

### 5.2) Commande optimale linéaire quadratique

Etant donné un système linéaire (non stationnaire)

$$\dot{x} = A(t)x + B(t)u, \quad x(t_0) = x_0, \quad (5.1)$$

on souhaite trouver la commande minimisant le critère quadratique

$$J = x'(t_1)Q_1x(t_1) + \int_{t_0}^{t_1} (x'Q(t)x + u'R(t)u) dt \quad (5.2)$$

où  $t_1$  est un instant final fixé,  $Q_1$ ,  $Q(t)$ ,  $R(t)$  sont des matrices (ou des fonctions matricielles) carrées symétriques  $n \times n$ ,  $n \times n$  et  $m \times m$  respectivement. On supposera toujours que  $R(t)$  est positive définie pour tout  $t$  (d'inverse bornée) et pour simplifier l'exposé, nous supposerons  $Q_1$  et  $Q(t)$  positives semi-définies.

La solution de ce problème peut s'exprimer en termes d'une nouvelle fonction matricielle  $n \times n$  symétrique positive semi-définie  $P(t)$  solution de l'équation différentielle suivante, dite équation de Riccati du problème:

$$\dot{P} + PA + A'P - PBR^{-1}B'P + Q = 0, \quad P(t_1) = Q_1. \quad (5.3)$$

elle est donnée par un feedback linéaire indépendant de  $x_0$  :

$$u^*(t) = -R^{-1}(t)B'(t)P(t)x(t). \quad (5.4)$$

Plus précisément, on a le résultat suivant :

**THÉORÈME 5.1.** L'équation de Riccati (5.3) admet une solution  $P(t)$  sur tout intervalle  $[t_0, t_1]$ , positive semi-définie, et  $P(t)$  est positive définie si la paire  $(Q^{\frac{1}{2}}, A)$  est complètement observable sur  $[t, t_1]$ (1). La solution du problème de minimisation de (5.2) sous la contrainte (5.3) est unique, et satisfait (5.4). La valeur optimale de  $J$  est  $J(u^*) = x_0'P(t_0)x_0$ .

On trouvera dans des ouvrages plus complets l'extension de ce résultat au cas où le critère comprend des termes "croisés" en  $x'Su$ , et des termes linéaires en  $u$  et  $x$ . De même, la dynamique peut contenir un terme additif  $v(t)$  connu.

On remarque que la solution prend la forme d'un "feedback d'état" (ou commandeur) non stationnaire de gain  $F = -R^{-1}B'P$ .

On verra au paragraphe 6.2 un équivalent en temps discret de cette théorie.

L'usage classique de ce résultat est le suivant. Le modèle (5.1) peut représenter un modèle linéarisé autour d'une trajectoire nominale qu'on désire suivre. Ainsi, les  $x_i$  représentent les écarts avec la trajectoire nominale, qu'on désire garder "petits". Le critère  $J$  représente une somme pondérée des carrés des écarts, les matrices  $Q(t)$  et  $R(t)$ , souvent choisies diagonales, incorporant les pondérations.  $Q_1$ , qui peut (de même que  $Q(t)$ ) être choisie nulle, permet de mettre une pondération plus forte sur un objectif final à atteindre. (cas typique : atterrissage d'un avion). Le concepteur peut avoir une idée de la façon dont il veut pondérer les différents écarts. Mais même s'il n'en n'a pas, cette méthode reste efficace. Il prendra d'abord une pondération arbitraire (disons des matrices identités partout), calculera  $P(t)$ , et simulera la réponse du système "optimal".

$$\dot{x} = (A - BR^{-1}B'P)x \quad (5.5)$$

pour différents états initiaux, et face à des perturbations diverses. Si une des variables,  $x_1$  par exemple, s'éloigne trop de la nominale, il faut augmenter son poids dans le critère en augmentant le coefficient  $q_{11}$  de  $Q$ , etc. Avec un peu d'expérience, cette procédure d'essais et d'erreurs permet de converger très rapidement vers un réglage satisfaisant.

Notons pour terminer que la dernière affirmation du théorème 5.1 donne une interprétation de la matrice  $P(t)$

### 5.3) Filtre de Kalman et principe de séparation

Nous considérons un système linéaire (non stationnaire) "excité" par des perturbations  $v(t)$  et  $w(t)$  :

$$\dot{x} = A(t)x + B(t)u + v(t), \quad x(t_0) = x_0, \quad (5.6)$$

$$y = C(t)x + w(t), \quad (5.7)$$

---

(1) Il s'agit ici du concept d'observabilité non stationnaire. Cette condition sera toujours remplie si  $A$  et  $Q$  sont des matrices constantes, avec  $(Q^{\frac{1}{2}}, A)$  complètement observable.

Dans ce paragraphe, on suppose que  $v(t)$  et  $w(t)$  sont des “bruits blancs gaussiens”. Une théorie rigoureuse demande malheureusement des outils de probabilité complexes (cette difficulté disparaît pour le modèle en temps discret, dont nous disons un mot in fine). Disons simplement que  $v(t)$  et  $w(t)$  sont des variables gaussiennes, centrées, et décorréllées dans le temps, au sens où

$$\mathbb{E}(v(t)v'(\tau)) = M(t)\delta(t - \tau), \quad (5.8.a)$$

$$\mathbb{E}(w(t)w'(\tau)) = N(t)\delta(t - \tau), \quad (5.8b)$$

$$\mathbb{E}(v(t)w'(\tau)) = 0. \quad (5.8c)$$

( $\delta$  est l’impulsion de Dirac. Dans le cas stationnaire, ce sont des processus de “densité spectrale”  $M$  et  $N$ , constantes pour toutes les fréquences). L’état initial  $x_0$  est lui-même une variable aléatoire gaussienne, indépendante de  $v$  et  $w$ , d’espérance  $\hat{x}_0$  et de covariance  $\Sigma_0$  données.

Le problème d’estimation optimale est d’évaluer

$$\hat{x}(t) = \mathbb{E}(x(t) \mid y(\tau), \tau \in [t_0, t]).$$

la solution de ce problème fait intervenir la matrice de covariance

$$\mathbb{E}(x(t) - \hat{x}(t))(x'(t) - \hat{x}'(t)) = \Sigma(t) \quad (5.9)$$

qui satisfait l’équation de Riccati suivante :

$$\dot{\Sigma} = A\Sigma + \Sigma A' - \Sigma C' N^{-1} C \Sigma + M, \quad \Sigma(t_0) = \Sigma_0, \quad (5.10)$$

et est donnée par la célèbre formule dite du filtre de Kalman :

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + B(t)u(t) + \Sigma(t)C'(t)N^{-1}(t)(y(t) - C(t)\hat{x}(t)). \quad (5.11)$$

Nous n’aurons pas le loisir de nous étendre ici sur l’étonnante dualité entre les formules (5.10)(5.11) d’une part, et (5.3)(5.5) d’autre part. Elle permet de déduire du théorème 5.1 l’existence d’une solution à l’équation de Riccati (5.10);

**THÉOREME 5.2.** L’équation de Riccati (5.10) admet une solution  $\Sigma(t)$  positive semi-définie sur tout intervalle  $[t_0, t_1]$ , et  $\Sigma(t)$  est positive définie si la paire  $(A, M^{\frac{1}{2}})$  est complètement commandable sur  $[t, t_1]$  (voir note dans théorème 5.1). La solution du problème d’estimation optimale est donnée par le filtre de Kalman (5.11). La matrice  $\Sigma(t)$  est la covariance de l’erreur d’estimation (5.9).

On remarquera que (5.11) a exactement la forme d’un “observateur” (non stationnaire) de gain  $K = \Sigma C' N^{-1}$  (remarquer la dualité avec le commandeur du paragraphe précédent). Ceci explique que dans le langage courant, tout compensateur de la forme “observateur commandeur” soit appelé “compensateur par filtre de Kalman”. On trouvera dans des ouvrages plus complets l’extension au cas où  $v$  et  $w$  ne sont pas décorréllés et où ils ont des espérances non nulles.

Ici, l’usage habituel de cette théorie est bien d’estimer les grandeurs relatives des incertitudes portant sur la dynamique et de celles portant sur la mesure. Alors, le filtre de Kalman permet de pondérer l’information “a priori” (5.6) avec les mesures (5.7) effectuées en fonctionnement.

Dans le cas où on a adopté un modèle en temps discret, correspondant à une commande numérique, constante par morceaux, on a

$$x(t+1) = A(t)x(t) + B(t)u(t) + v(t), \quad (5.6d)$$

$$y(t) = C(t)x(t) + w(t). \quad (5.7d)$$

Ici  $v(t)$  et  $w(t)$  sont simplement des suites décorréllées. Le filtre s’écrit

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) + A\Sigma(t)C'(N + C\Sigma(t)C')^{-1}(y(t) - C\hat{x}(t)), \quad (5.11d)$$

avec

$$\Sigma(t+1) = A\Sigma(t)A' + A\Sigma(t)C'(N + C\Sigma(t)C')^{-1}C\Sigma(t)A' + M, \quad \Sigma(t_0) = \Sigma_0. \quad (5.10d)$$

L'équation de Riccati a une forme plus complexe, mais le filtre de Kalman garde la forme d'un observateur "de Luenberger" et donc partage son caractère de prédicteur, permettant l'exploitation effective en " temps réel" des idées qui suivent

L'idée naturelle pour commander un système présentant des perturbations et avec des mesures imparfaites est d'associer les deux techniques ci-dessus : définir un critère quadratique de la forme (5.2), calculer le gain de feedback optimal par (5.3) et (5.4), et l'appliquer à l'estimée  $\hat{x}$  calculée par (5.10),(5.11). Que ceci associe l'efficacité des deux méthodes sans effet pervers est garanti par le grand résultat suivant, dit principe de séparation.

**THÉORÈME 5.3.** La procédure décrite ci-dessus conduit à la commande qui minimise l'espérance de  $J$  parmi toutes celles qui dépendent de  $y(t)$  de façon causale.

Ainsi, la procédure de l'observateur-commandeur peut être fondée soit sur des arguments d'optimalité, soit sur des arguments de "placement de pôles". En fait, dans le cas stationnaire, ces deux procédures sont voisines, et dans la pratique on mélange parfois les deux considérations.

#### 5.4) Le cas stationnaire

En l'état, la théorie ci-dessus présente l'inconvénient que même si le modèle  $(C, A, B)$  est stationnaire, les matrices  $P$  et  $\Sigma$  sont variables, et donc aussi les gains du filtre et du commandeur. Ceci vient de ce que le problème a été posé sur un horizon fini. On peut retrouver une théorie stationnaire de la façon suivante. supposons que  $(Q^{\frac{1}{2}}, A, B)$  est stationnaire canonique, soit  $Q \geq 0$  et  $R > 0$ . Posons

$$J = \int_0^{\infty} (x'Qx + u'Ru) dt ,$$

et imposons de surcroit que la commande retenue stabilise le système.

**THÉORÈME 5.4.** Sous les hypothèses ci-dessus, l'équation de Riccati (5.3) intégrée depuis  $t_1 = 0$  avec  $P(0) = 0$ , a une solution  $P(t)$  qui tend, quand  $t$  tends vers  $-\infty$ , vers une matrice symétrique positive  $P^*$  solution maximum, et unique solution positive définie, de l'équation de Riccati algébrique

$$PA + A'P - PBR^{-1}B'P + Q = 0,$$

La commande stabilisante qui minimise  $J$  est donnée par la formule (5.4) où  $P(t)$  est remplacée par  $P^*$ .

Ce théorème permet d'appliquer la procédure du paragraphe 5.2 pour concevoir un régulateur stationnaire. Il faut faire attention que l'équation de Riccati algébrique a, d'habitude, de nombreuses solutions. Le théorème indique comment calculer celle qui nous intéresse comme valeur limite de la solution d'une équation différentielle intégrée en temps rétrograde.

Dualement, pour un système de la forme (5.6)(5.7) stationnaire  $(A, B, C, M$  et  $N$  constantes), on peut supposer que les mesures  $y(t)$  sont disponibles depuis très longtemps, et chercher  $\hat{x}(t) = E(x(t) | y(\tau), \tau \in (-\infty, t))$ . Alors, supposons que  $(C, A, M^{\frac{1}{2}})$  est canonique. On a le théorème dual du précédent.

**THÉORÈME 5.5.** Sous les hypothèses ci-dessus, l'équation de Riccati (5.10) intégrée depuis  $\Sigma(0) = 0$  a une solution qui converge, quand  $t$  tend vers l'infini, vers une matrice symétrique  $\Sigma^*$ , unique solution positive définie de l'équation de Riccati algébrique associée à (5.10). La solution du problème d'estimation stationnaire est donnée par le filtre de Kalman (5.11) où  $\Sigma(t)$  est remplacée par  $\Sigma^*$ .

On a donc le moyen de concevoir un observateur-commandeur stationnaire. Naturellement, le principe de séparation subsiste dans cette version de la théorie.

Il est intéressant de faire un lien entre cette théorie et la théorie "modale". Nous donnons un bref aperçu de la théorie algébrique de la commande optimale qui fait, entre autres, ce lien.

Au problème (5.1)(5.2) stationnaire, associons sa matrice hamiltonienne  $\mathcal{H}$  qui est une matrice  $2n \times 2n$  définie par blocs de la façon suivante

$$\mathcal{H} = \begin{pmatrix} A & -BR^{-1}B' \\ -Q & -A' \end{pmatrix} .$$

On montre facilement que si  $\lambda$  est valeur propre de cette matrice,  $-\lambda$  l'est aussi. On montre aussi (moins facilement) que sous les hypothèses de canonicité du théorème 5.4, elle n'a pas de valeur propre imaginaire pure. Ainsi, elle a  $n$  valeurs propres à parties réelles négatives (stables) et leur  $n$  opposées.

On a un premier résultat spectral :

**THÉOREME 5.6.** Les valeurs propres de la matrice dynamique “bouclée”  $A - BR^{-1}B'P^*$  sont les valeurs propres stables de  $\mathcal{H}$ .

On peut utiliser cette théorie pour calculer simplement  $P^*$  à partir des vecteurs propres correspondant à ces valeurs propres.

On peut donner des résultats de nature “fréquentielle”. Voici l’un d’eux. Considérons la matrice rationnelle  $m \times m$  (toujours avec les mêmes hypothèses)

$$\Gamma(s) = R + B'(-sI - A')^{-1}Q(sI - A)^{-1}B$$

**THÉOREME 5.7.** Les valeurs propres de  $A - BR^{-1}B'P^*$  sont les racines à partie réelle négative du déterminant de  $\Gamma(s)$

## 5.5) Extensions diverses

### 5.5.1) Systèmes gouvernés par des équations aux dérivées partielles

Nous quittons momentanément le monde des systèmes “en dimension finie”, pour donner un bref aperçu d’une extension très importante de cette théorie aux systèmes gouvernés par des équations aux dérivées partielles. Nous nous limiterons aux systèmes paraboliques du deuxième ordre.

Imaginons un volume  $\Omega$  de l’espace physique (un four) de paroi  $\Gamma$  adiabatique, chauffé par apport de chaleur en certains points de  $\Omega$  (des brûleurs). Soit  $x \in \Omega$  la variable “d’espace” parcourant  $\Omega$ , et  $y(x, t)$  la température au point  $x$  à l’instant  $t$ . Un modèle d’évolution de  $y$  est

$$\frac{\partial y}{\partial t} = - \sum_{i=1}^3 \frac{\partial^2 y}{\partial x_i^2} + B(x)u(t),$$

$$\forall x \in \Gamma, \quad \forall t, \quad \frac{\partial y}{\partial n}(x, t) = 0.$$

$\partial y / \partial n$  désigne la dérivée normale sur  $\Gamma$ .  $B(x)$  répartit le flux de chaleur  $u(t)$  sur les régions où il est appliqué. Ceci est un cas particulier d’un système de la forme

$$\dot{y} = Ay + Bu$$

où  $y(t)$  est une “variable répartie”  $y(x, t)$ , et  $A$  l’opposé d’un opérateur elliptique, muni d’une condition au bord, ici de Neuman, mais qui pourrait aussi être de Dirichlet (i.e.  $y(x, t)$  donné pour tout  $x \in \Gamma$ , et tout  $t$ ) ou mixte : d’un type sur une partie de la frontière, de l’autre ailleurs.

On peut à nouveau doter le système d’un critère quadratique, de la forme

$$J = (y(t_1), Q_1 y(t_1)) + \int_{t_0}^{t_1} [(y(t), Q(t)y(t)) + R(t)u^2(t)] dt,$$

où la notation  $(y, z)$  désigne le produit scalaire dans  $L^2(\Omega)$  :

$$(y, z) = \int_{\Omega} y(x)z(x) dx.$$

On peut encore donner un sens à l’équation de Riccati, et obtenir un théorème analogue à 5.1. Mais ce n’est plus un moyen efficace de calcul. On doit, en pratique, se contenter de solutions “en boucle ouverte”, donc dépendant de l’état initial, obtenues par des méthodes d’analyse numérique exploitant la technique de l’état adjoint (cf chapitre 6).

### 5.5.2) Horizon glissant

Il est clair que puisque

$$\int_{t_0}^{t_1} L(\tau) d\tau = \int_{t_0}^t L(\tau) d\tau + \int_t^{t_1} L(\tau) d\tau,$$

on peut, à l'instant  $t$ , ignorer la dépense déjà faite, et, connaissant  $x(t)$  ne minimiser que la partie restant à couvrir du coût

On a vu ci-dessus que, pour obtenir un régulateur stationnaire, il pouvait être utile de prendre  $t_1 = \infty$ , ce qui fait que la commande appliquée à chaque instant  $t$  minimise

$$\int_t^{\infty} (x'Qx + u'Ru) dt.$$

Il a été proposé, au lieu d'un horizon infini, de se contenter d'un horizon "suffisamment long"  $T$ , conduisant à appliquer à l'instant  $t$  la valeur à cet instant de la commande qui minimise

$$\int_t^{t+T} (x'Qx + u'Ru) dt.$$

C'est ce qu'on appelle la technique de "l'horizon glissant". On voit que, sous cette forme (et dans le cas stationnaire), elle consiste à n'intégrer l'équation de Riccati que sur l'intervalle  $[-T, 0]$  au lieu de  $(-\infty, 0]$ , ou plutôt au lieu d'attendre d'avoir atteint un état stationnaire.

On peut raffiner cette idée en définissant une "courbe de raccordement" souhaitée  $x_d(t + \tau), u_d(t + \tau)$  reliant  $x(t)$  à 0, et en minimisant

$$\int_t^{t+T} [(x - x_d)'Q(x - x_d) + (u - u_d)'R(u - u_d)] dt.$$

Les auteurs ont aussi discuté le choix de la longueur de l'horizon de prédiction  $T$ , et, surtout dans le cas discret, la longueur de l'intervalle sur lequel la commande est appliquée avant d'en calculer une autre. L'ensemble de ces idées et techniques est connu sous le nom de "commande prédictive" ou "Model Reference Predictive Control".

### 5.5.3) Commande adaptative

La commande adaptative a été (rapidement) présentée au chapitre précédent. Nous voulons seulement montrer ici l'usage du filtre de Kalman pour identifier un processus, et adapter le contrôleur.

Considérons un modèle monovarié ( $u$  et  $y$  scalaires) dont la forme ARMA s'écrit, avec un bruit blanc  $w(t)$ , sous la forme

$$y(t) + a_1y(t-1) + \dots + a_p y(t-p) = b_1u(t-1) + \dots + b_q u(t-q) + w(t), \quad (5.13)$$

où on supposera  $b_1 \neq 0$ . On montre que la commande qui minimise la variance du processus  $y(\cdot)$  est

$$u(t) = \frac{1}{b_1} [a_1y(t) + a_2y(t-1) + \dots + a_p y(t-p+1) - b_2u(t-1) - \dots - b_q u(t-q+1)]. \quad (5.14)$$

(c'est à dire la commande qui donnerait  $y(t+1) = 0$  si  $w(t+1) = 0$ ). On remarquera que ce régulateur ne peut être utilisé tel quel que si le système d'origine est à "phase minimale", c'est à dire si le polynôme  $b_1z^{q-1} + \dots + b_q$  a des racines de module inférieur à l'unité. Autrement (5.14) est un système dynamique en  $u(t)$  instable. Les régulateurs fondés sur la théorie précédente sont beaucoup moins sensibles à ce problème. Mais nous nous intéressons au cas où les coefficients  $a_i$  et  $b_j$  sont mal connus, voire susceptibles de varier lentement.

Appelons  $\gamma(t)$  la matrice ligne

$$\gamma(t) = [-y(t-1) \quad -y(t-2) \quad \dots \quad -y(t-p) \quad u(t-1) \quad \dots \quad u(t-q)]$$

et  $\theta(t)$  le vecteur inconnu des coefficients :

$$\theta(t) = [a_1 \quad a_2 \quad \dots \quad a_p \quad b_1 \quad b_2 \quad \dots \quad b_q]'$$

le modèle (5.13) s'écrit aussi

$$y(t) = \gamma(t)\theta(t) + w(t). \quad (5.15)$$

Maintenant, l'hypothèse que le vecteur  $\theta$  est constant, ou varie lentement, peut s'écrire

$$\theta(t+1) = \theta(t) + v(t), \quad (5.16)$$

où  $v(t)$  sera un bruit blanc centré de covariance nulle si  $\theta$  est constant, ou dont on choisira la covariance plus ou moins grande pour exprimer nos hypothèses sur la vitesse de variation possible de  $\theta$ .

On remarquera alors que les équations (5.16)(5.15) ont exactement la forme discrète correspondant à (5.6d)(5.7d), avec  $\theta$  pour état. En effet, à l'instant  $t$ ,  $\gamma(t)$  est connu. On peut donc, à condition de propager  $\Sigma(t)$  par (5.10d) en temps réel, évaluer l'estimée  $\hat{\theta}$  de  $\theta$  à l'aide d'un filtre de Kalman discret.

Une méthode de commande adaptative naturelle est alors d'appliquer la formule (5.14) en y remplaçant les  $a_i$  et  $b_j$  par leurs estimées  $\hat{a}_i$  et  $\hat{b}_j$  issues de  $\hat{\theta}$ . C'est là la version la plus élémentaire de régulateur autoajustable (self tuning regulator). Cette méthode a été étendue aux systèmes multivariables. Bien que sous sa forme simple elle présente quelques gros défauts théoriques (elle peut ne pas converger, ou converger vers une mauvaise estimée) elle a été utilisée avec succès pour la régulation de processus industriels.

### 5.6) La commande $H_\infty$ comme problème linéaire quadratique.

Ceci constitue une nouvelle approche du problème évoqué au paragraphe 3.6 (1)

Nous reformulons légèrement le problème de commande d'un système linéaire perturbé tel que présenté au paragraphe 5.3. Le système peut s'écrire

$$\dot{x} = Ax + Bu + Ev, \quad (5.17)$$

$$y = Cx + Dv, \quad (5.18)$$

où on adopte pour  $v(\cdot)$  le modèle d'un bruit blanc unitaire (de covariance  $I\delta(t-\tau)$ ) et où

$$EE' = M, \quad DD' = N, \quad ED' = 0. \quad (5.19)$$

Comme précédemment, on supposera toujours  $N$  inversible. La restriction  $ED' = 0$  n'est nécessaire que pour nous limiter au cas (5.8c), ce qui simplifie les équations tant dans la solution du problème L.Q.G. du paragraphe 5.3, que dans celle du problème  $H_\infty$  présenté ici.

L'objectif de commande L.Q.G. peut être décrit comme : assurer que  $J$  (5.2) soit petit en présence de perturbations inconnues. Et on s'est donné un modèle stochastique des perturbations. Ici, on va chercher à atteindre le même objectif (vague) sans se donner de modèle de  $v(\cdot)$ , mais en la supposant de carré sommable (ce qui n'est pas le cas du bruit blanc). Pour un régulateur donné, nous écrirons

$$\gamma^2 = \sup_{v \in L^2} \frac{J}{\|v\|^2}, \quad (5.20)$$

et nous appellerons  $\gamma$  son "facteur d'atténuation". En effet, la définition (5.20) garantit que

$$J \leq \gamma^2 \|v\|^2. \quad (5.21)$$

La formulation précise de notre objectif sera donc de minimiser le facteur d'atténuation. On notera  $\gamma^*$  le facteur d'atténuation optimal. Par souci de simplicité nous nous restreindrons au cas  $x_0 = 0$ . (Sinon il faut, comme en commande stochastique, inclure  $x_0$  dans les perturbations).

Montrons en quoi le problème posé ici se rattache au problème de commande  $H_\infty$  présenté au paragraphe 3.6. Pour nous restreindre au cas linéaire, nous imposons que le contrôleur  $u = K(y)$  soit linéaire. Ignorons aussi le terme en  $x(t_1)$  dans  $J$ . Alors, pour des matrices  $H$  et  $G$  bien choisies, on peut poser  $z = Hx + Gu$ , et avoir  $J = \|z\|^2$ , norme au sens  $L^2$ . Le régulateur  $K$  étant linéaire, il existe un opérateur linéaire  $T_K$  tel que  $z = T_K v$ . Alors,  $\gamma$  défini par (5.20) est la norme d'opérateur de  $T_K$ . Si enfin on se met dans le

---

(1) Cette approche de la commande  $H_\infty$  date de 1989, et de 1990 pour les cas en mesures imparfaites

cas stationnaire, (cf paragraphe 5.4), en exigeant que le système soit observable par  $z$  pour garantir qu'un contrôleur rendant  $J$  fini le stabilise, alors  $T_K$  peut être représenté par sa fonction de transfert. Et sa norme d'opérateur de  $L^2$  dans  $L^2$  est la norme  $H_\infty$  de sa fonction de transfert. Donc dans ce cas particulier le problème posé ici coïncide avec un problème  $H_\infty$ , dont on vérifie par ailleurs facilement que la plupart des problèmes qui ont été considérés sous l'angle  $H_\infty$  (dont la poursuite de modèle évoquée au chapitre 3) s'y ramènent. Mais la formulation présente est plus générale.

Comme dans toute la théorie  $H_\infty$ , ce qui est facile est, étant donné  $\gamma$ , vérifier s'il existe un contrôleur  $K$  satisfaisant (5.21), et si oui en trouver un. La façon d'aborder le problème de commande  $H_\infty$  sera donc d'essayer des coefficients d'atténuation de plus en plus petits, jusqu'à ce que le test d'existence échoue.

Montrons maintenant que trouver un contrôleur  $K$  qui garantisse (5.21) revient à résoudre un problème de min-max, En effet, (5.21) peut aussi s'écrire

$$\forall v \in L^2, \quad \|z\|^2 - \gamma^2 \|v\|^2 \leq 0$$

soit encore

$$\sup_{v \in L^2} (\|z\|^2 - \gamma^2 \|v\|^2) \leq 0, \quad (5.22)$$

et un contrôleur  $K$  existe qui satisfait cette condition si et seulement si

$$\inf_K \sup_{v \in L^2} (\|z\|^2 - \gamma^2 \|v\|^2) \leq 0. \quad (5.23)$$

Remarquons que par linéarité, à  $K$  fixé, le sup dans (5.22) est soit négatif ou nul (en fait nul) soit infini. Donc si  $\gamma < \gamma^*$ , l'inf sup dans (5.23) est infini, et si  $\gamma > \gamma^*$  il est fini (et nul).

Le problème de déterminer l'inf sup est un jeu différentiel (voir chapitre 6). Nous donnerons d'abord sa solution dans le cas de l'information parfaite, c'est à dire où l'état  $x$  du système peut être mesuré directement. A nouveau par souci de simplicité, nous nous restreindrons au cas où  $J$  s'écrit comme en (5.2), c'est à dire où  $H'H = Q$ ,  $G'G = R$ , et  $H'G = 0$ . Lever cette dernière restriction ne fait que compliquer un peu les équations.

La solution du jeu (5.23) s'exprime en fonction d'une matrice symétrique  $P(t)$  solution de l'équation de Riccati du jeu :

$$\dot{P} + PA + A'P - P(BR^{-1}B' - \gamma^{-2}M)P + Q = 0, \quad P(t_1) = 0$$

et la formule du contrôleur optimal est la même qu'au paragraphe 5.2

$$u(t) = -R^{-1}(t)B'(t)P(t)x(t). \quad (5.25)$$

Mais maintenant l'équation de Riccati peut n'avoir pas de solution sur  $[t_0, t_1]$ , sa solution intégrée depuis  $t_1$  en temps rétrograde divergeant vers  $+\infty$  avant que  $t$  ait atteint  $t_0$ . On dit que le problème présente un *point conjugué*. Dans ce cas l'inf sup de (5.23) est infini. On peut énoncer le théorème suivant:

**THÉORÈME 5.8.** Si  $\gamma > \gamma^*$ , l'équation de Riccati (5.25) a une solution sur  $[t_0, t_1]$ , (positive semi-définie), et un contrôleur qui assure (5.21) est donné par (5.25). Si  $\gamma < \gamma^*$ , l'équation (5.24) n'a pas de solution définie sur tout l'intervalle  $[t_0, t_1]$ .

Ce résultat s'étend au cas stationnaire, qui fera intervenir l'équation de Riccati algébrique

$$PA + A'P - P(BR^{-1}B' - \gamma^{-2}M)P + Q = 0, \quad (5.26)$$

de la façon suivante :

**THÉORÈME 5.9.** La paire  $(H, A)$  est supposée complètement observable. (1) Si  $\gamma > \gamma^*$ , l'équation de Riccati algébrique (5.26) a une solution positive définie  $P^*$ , et l'équation 5.24 intégrée en temps rétrograde depuis 0 a une solution qui tend vers  $P^*$  quand  $t$  tend vers  $-\infty$ . Dans ce cas, un contrôleur qui stabilise

---

(1) En fait, il ne peut exister  $P^*$ , et un contrôleur de norme  $H_\infty$  bornée, que si en outre  $(A, B)$  est complètement accessible.

le système et assure (5.21) est donné par (5.25) où on remplace  $P(t)$  par  $P^*$ . Si  $\gamma < \gamma^*$ , (5.26) n'a pas de solution positive définie, et la solution  $P(t)$  de (5.24) diverge pour  $t \rightarrow t^* > -\infty$ .

D'autres caractérisations de  $P^*$  existent, notamment en termes spectraux.

Donnons maintenant la solution de (5.23) dans le cas en information imparfaite, c'est à dire où seule est accessible une sortie du type (5.18). A nouveau par souci de simplicité, on adoptera les notations (5.19) et la restriction  $ED' = 0$  qu'elles contiennent.

La solution de ce jeu différentiel en information imparfaite fait intervenir un principe de séparation similaire à celui de problème *LQG* : nous allons déterminer une variable  $\hat{x}(t)$  (plus vraiment une estimée) qu'il faut mettre à la place de  $x(t)$  dans (5.25) pour obtenir un régulateur optimal. Le calcul de cette variable fait intervenir une équation de Riccati duale de la précédente :

$$\dot{\Sigma} = A\Sigma + \Sigma A' - \Sigma(C'N^{-1}C - \gamma^{-2}Q)\Sigma + M, \quad \Sigma(t_0) = 0 \quad (5.27)$$

et un nouveau filtre, ressemblant à un filtre de Kalman. Il s'exprime en fonction de la perturbation maximisante  $\hat{v} = \gamma^{-2}E'P\hat{x}$  comme

$$\dot{\hat{x}} = A\hat{x} + Bu * (\hat{x}) + E\hat{v} + (I - \gamma^{-2}\Sigma P)^{-1}\Sigma C'N^{-1}(y - C\hat{x} - D\hat{v}). \quad (5.28)$$

Le contrôleur optimal est donc

$$u * (\hat{x}) = -R^{-1}B'P\hat{x}. \quad (5.29)$$

Deux nouvelles façons d'échouer dans le calcul du contrôleur peuvent se présenter : soit l'équation de Riccati (5.27) diverge avant  $t_1$ , soit la matrice  $(I - \gamma^{-2}\Sigma P)$  cesse d'être inversible. L'une comme l'autre indique que  $\gamma \leq \gamma^*$ . On peut préciser ce résultat par le théorème suivant :

**THÉOREME 5.10.** Si  $\gamma > \gamma^*$ , les équations de Riccati (5.24) et (5.27) ont une solution sur  $[t_0, t_1]$ , et la matrice  $(I - \gamma^{-2}\Sigma P)$  est inversible sur cet intervalle. Un contrôleur optimal est donné par (5.28)(5.29). Si  $\gamma < \gamma^*$ , l'une des trois conditions ci-dessus n'est pas satisfaite.

On a enfin l'extension logique au problème stationnaire.

**THÉOREME 5.11.** Le système (H,A,E) est supposé stationnaire et minimal. (Complètement observable et accessible). Pour le problème stationnaire, si  $\gamma > \gamma^*$ , les équations de Riccati algébriques du problème admettent des solutions positives définies  $P^*$  et  $\Sigma^*$ , limites des solutions de (5.24) et (5.27) et  $\rho(\Sigma P) < \gamma^2$ , garantissant que la matrice  $(I - \gamma^{-2}\Sigma^*P^*)$  est inversible. Un contrôleur optimal est obtenu en remplaçant  $P(t)$  et  $\Sigma(t)$  par  $P^*$  et  $\Sigma^*$  dans (5.28) et (5.29). Si  $\gamma < \gamma^*$ , l'une de ces conditions n'est pas satisfaite.

Naturellement, le choix d'un critère  $J$  adapté au problème de régulation considéré peut être rapproché du choix discuté au paragraphe 5.2. Le choix d'un contrôleur  $H_\infty$  plutôt que gaussien est réputé être plus prudent, moins efficace bien sûr, si  $v$  est vraiment gaussien, mais moins sensible aux erreurs de modélisation, puisqu'aussi bien on n'a pas modélisé la perturbation.

On a aussi vu au chapitre 3, paragraphe 6, que savoir minimiser la norme  $H_\infty$  de la fonction de transfert du système bouclé avait d'autres applications intéressantes.

## 6) Commande optimale

### 6.1) La problématique

Nous avons résolu au paragraphe 5.2 un problème de commande optimale. Nous généraliserons ici à une forme non linéaire. Les problèmes considérés seront définis par les ingrédients suivants :

- Un système dynamique dans  $\mathbb{R}^n$  (c.a.d. où  $x$  a  $n$  coordonnées  $x_i, i = 1 \dots n$ )

$$\dot{x} = f(x, u, t), \quad x(t_0) = x_0. \quad (6.1)$$

On fait des hypothèses de régularité et de croissance adéquates sur  $f$  pour assurer l'existence et l'unicité de la trajectoire  $x(\cdot)$  pour toutes les commandes admissibles.

- Un ensemble de commandes admissibles, défini par un sous ensemble fermé  $U$  de  $\mathbb{R}^m$  (éventuellement  $\mathbb{R}^m$  tout entier). Les commandes admissibles seront les fonctions  $t \mapsto u(t)$  mesurables (ou plus simplement continues par morceaux) telles que pour (presque) tout  $t, u(t) \in U$ . On notera  $\Omega$  cet ensemble.

- Un objectif à atteindre, ou cible, sous-ensemble fermé  $C$  de  $\mathbb{R}^n \times \mathbb{R}$ . L'instant final  $t_1$  du problème sera le premier instant tel que

$$(x(t_1), t_1) \in C. \quad (6.2)$$

- Un critère à minimiser sur  $\Omega$ , donné par

$$J(u(\cdot)) = K(x(t_1), t_1) + \int_{t_0}^{t_1} L(x(t), u(t), t) dt. \quad (6.3)$$

A nouveau des hypothèses adéquates sont faites sur  $K$  et  $L$  pour garantir l'existence de  $J$  pour tout  $u(\cdot)$  qui satisfait (6.2) pour un  $t_1 < \infty$ .

Il faut discuter quelques cas particuliers de cette forme générale. Le terme  $K$ , dans (6.3), représente un coût lié au point atteint, et peut permettre de pénaliser un objectif plus ou moins bien atteint. Le terme  $L(x, u, t)$  représente la dépense associée au choix de la commande  $u$  à l'instant  $t$  en l'état  $x$ .

Le problème du paragraphe 5.2 est caractérisé, non seulement par une dynamique (6.1) linéaire et un critère (6.3) quadratique, mais aussi par un instant final  $t_1$  donné a priori. Ceci rentre dans le formalisme (6.2) en posant  $C = \mathbb{R}^n \times \{t_1\}$ ,  $t_1$  donné. Dans ce cas, la fonction  $K$  dans (6.3) n'a pas matière à dépendre de  $t_1$ .

Un autre cas classique est celui où  $C = \{0\} \times [t_0, +\infty)$ , c'est à dire que  $t_1$  est le premier instant tel que  $x(t_1) = 0$ , et  $K = 0$ ,  $L = 1$ . Soit  $J = t_1 - t_0$ . C'est le plus classique des problèmes à temps minimum. On peut naturellement l'étendre à une cible  $C_0$  autre que l'origine, voire une cible mobile avec le temps  $C_t$ . On doit alors prendre  $C = \{(x, t) \mid x \in C_t\}$ .

Enfin, il est utile de considérer la version discrète de ce problème

$$x_{t+1} = f_t(x_t, u_t), \quad x_{t_0} = x_0, \quad (6.4a)$$

$$J(u) = K(x_{t_1}, t_1) + \sum_{t=t_0}^{t_1-1} L_t(x_t, u_t). \quad (6.4b)$$

## 6.2) Programmation dynamique et équation de Hamilton Jacobi

Une approche très simple pour résoudre le problème (6.4) est celle de la programmation dynamique. Son principal inconvénient est qu'elle n'est praticable qu'à condition que la dimension des espaces d'état et de commande soient petites.

Nous disons que le problème (6.4) "commence en  $(x_0, t_0)$ ". Considérons alors tous les problèmes de la forme (6.4), mais "commençant en  $(\xi, \tau)$ ", pour  $\xi$  quelconque dans  $\mathbb{R}^n$ , et  $\tau > t_0$ . C'est à dire qu'on aura  $x(\tau) = \xi$ , et

$$J_{\xi, \tau}(u) = K(x_{t_1}, t_1) + \sum_{t=\tau}^{t_1-1} L_t(x_t, u_t). \quad (6.5)$$

et posons,

$$V(\xi, \tau) = \min_{u \in \Omega} J_{\xi, \tau}(u(\cdot)). \quad (6.6)$$

Supposons qu'à l'instant  $t$  on soit en  $x_t = \bar{x}$ , et qu'on décide d'appliquer la commande  $u$  qui nous envoie à l'instant  $t+1$  en  $x_{t+1} = f_t(\bar{x}, u)$ . Alors, de  $t+1$  jusqu'à la fin, on ne devrait pas "dépenser" plus que  $V(x_{t+1}, t+1)$ , menant à une dépense au mieux dans ce cas de

$$J_{\bar{x}, t} = L_t(\bar{x}, u) + V(f_t(x, u), t+1)$$

On devrait donc choisir  $u$  minimisant ce coût, d'où (en remplaçant  $\bar{x}$  par  $x$ )

$$V(x, t) = \min_{u \in U} [L_t(x, u) + V(f_t(x, u), t+1)]. \quad (6.7)$$

C'est l'équation de Bellman, qui permet de calculer  $V$ , et la commande optimale, si on sait l'initialiser. Or on se convainc facilement que  $V(x_{t_1}, t_1) = K(x_{t_1}, t_1)$ , soit

$$\forall (x, t) \in C, \quad V(x, t) = K(x, t). \quad (6.8)$$

La méthode de solution consiste à discrétiser aussi l'état. On initialise  $V$  sur la cible suivant (6.8), puis on applique (6.7) pour calculer  $V$  à tous les états d'où la cible est atteinte en un pas de temps, s'il le faut en "essayant" toutes les valeurs permises de  $u$ , et en retenant la meilleure. On "recule" alors d'un pas de temps, connaissant les valeurs de  $V$  aux points qui sont atteints dans (6.7), et ainsi de suite.

Quand tout le tableau en  $(x \times t)$  est calculé, on a non seulement  $V$ , et donc  $V(x_0, t_0)$  qui est la valeur optimale du critère, mais surtout la commande optimale  $u^*$  à appliquer en tout point  $(x, t)$ . On a donc déterminé le feedback optimal  $u = \phi^*(x, t)$ . Celui-ci pourra être appliqué même si des perturbations, non modélisées ici, font sortir l'état de la trajectoire optimale qu'on aurait calculée.

Un cas particulier important est celui du problème linéaire quadratique :

$$x_{t+1} = A_t x_t + B_t u_t,$$

$$J(u(\cdot)) = \sum_{t=t_0}^{t_1-1} [x'_{t+1} Q_{t+1} x_{t+1} + u'_t R_t u_t].$$

On trouve alors une solution "analytique" de (6.7)(6.8) en posant

$$V(x, t) = x' P_t x$$

la matrice symétrique  $P$  satisfaisant l'équation de Riccati discrète

$$P_t = A' P_{t+1} A - A' P_{t+1} B (R + B' P_{t+1} B)^{-1} B' P_{t+1} A + Q, \quad P_{t_1} = Q_{t_1},$$

(les matrices  $A$ ,  $B$ ,  $Q$  et  $R$  ci-dessus sont évaluées en  $t$ ) avec

$$u_t = -(R + B' P_{t+1} B)^{-1} B' P_{t+1} A x.$$

Ceci est l'équivalent discret du résultat du paragraphe 5.2.

Si le modèle discret (6.4) était obtenu par une discrétisation en temps d'un modèle continu, il est tentant de faire tendre le pas de discrétisation vers zéro dans l'équation de Bellman. Ceci mène formellement à l'équation de Hamilton Jacobi Carathéodory :

$$\frac{\partial V}{\partial t} + \min_{u \in U} \left( L(x, u, t) + \frac{\partial V}{\partial x} f(x, u, t) \right) = 0. \quad (6.9)$$

les équations (6.8)(6.9) permettent d'énoncer une condition suffisante :

**THÉOREME 6.1.** S'il existe une solution continûment dérivable de (6.8)(6.9), et si la minimum en (6.9) est obtenu en  $u = \phi^*(x, t)$ , feedback admissible, alors  $\phi^*$  est un feedback optimal pour tous les états initiaux tels que la trajectoire engendrée coupe la cible.

L'intérêt de ce théorème est surtout théorique, sauf dans le cas linéaire quadratique où il permet d'arriver rapidement aux formules du paragraphe 5.2.

### 6.3) Méthodes variationnelles et principe du maximum de Pontryaguine

La plupart des méthodes numériques permettant de calculer une solution au problème (6.1) à (6.3) sont fondées sur une analyse variationnelle, c'est à dire une expression "au premier ordre" des petites variations de  $J$  pour de petites variations de  $u(\cdot)$ , et de l'instant final  $t_1$ .

En effet, dans la suite, nous considérons  $J$  comme dépendant des variables  $u$  et  $t_1$ , dans  $\Omega \times \mathbb{R}$ . Nous supposons que la cible  $C$  est donnée par une fonction  $\gamma$  de  $\mathbb{R}^n \times \mathbb{R}$  dans  $\mathbb{R}^r$ :

$$C = \{x, t \mid \gamma(x, t) = 0\},$$

de sorte que les variables  $u(\cdot)$  et  $t_1$  sont liées par les  $r$  contraintes

$$\gamma_i(x(t_1), t_1) = 0, \quad i = 1 \dots r. \quad (6.10)$$

$x(t_1)$  dépendant bien sûr de  $u(\cdot)$  au travers de (6.1).

Introduisons, pour simplifier les notations, une nouvelle fonction scalaire  $H$  de  $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$  dans  $\mathbb{R}$  définie par

$$H(x, \lambda, u, t) = L(x, u, t) + \lambda' f(x, u, t) = L(x, u, t) + \sum_{k=1}^n \lambda_k f_k(x, u, t). \quad (6.11)$$

Supposons  $u(\cdot)$  et  $t_1$  fixés. On peut calculer  $x(t)$  numériquement en intégrant (6.1) à l'aide d'un programme d'intégration d'équation différentielle. On peut calculer l'état adjoint  $\lambda(t) = (\lambda_1(t), \lambda_2(t), \dots, \lambda_n(t))$  défini par

$$\lambda_k(t_1) = \frac{\partial K}{\partial x_k}(x(t_1), t_1), \quad \dot{\lambda}_k(t) = -\frac{\partial H}{\partial x_k}(x(t), \lambda(t), u(t), t). \quad (6.12)$$

Le résultat central, connu sous le nom de formule de l'état adjoint, est le suivant : sous des hypothèses de régularité adéquates, on a

$$J(u(\cdot) + \delta u(\cdot), t_1 + dt_1) - J(u(\cdot), t_1) = \int_{t_0}^{t_1} \sum_k \frac{\partial H}{\partial u_k}(t) \delta u_k(t) dt + \left( H(t_1) + \frac{\partial K}{\partial t_1} \right) dt_1 + 0(\delta u) + 0(dt_1), \quad (6.13)$$

où  $0(\delta u)$  et  $0(dt_1)$  tendent vers zéro plus vite que  $\|\delta u\|$  et  $|dt_1|$  respectivement. C'est à dire que la fonction vectorielle  $\partial H / \partial u(t)$  joue le rôle de gradient de  $J$  par rapport à  $u(\cdot)$ , et  $H(t_1) + \partial K / \partial t_1$  le rôle de dérivée par rapport à  $t_1$ .

Remarquons en outre que pour chaque  $i = 1, \dots, p$ , la fonction  $\gamma_i(x(t_1), t_1)$  dépend de  $u(\cdot)$  et  $t_1$  au même titre que  $K(x(t_1), t_1)$ . On peut donc calculer ses dérivées en adoptant la même méthode, où  $K$  est remplacé par  $\gamma_i$  et où  $L$  est nulle. Ceci mène à introduire un (autre) état adjoint  $\mu_i(t)$  par coordonnée de  $\gamma$ , et permet de calculer leurs dérivées.

Ainsi toutes les méthodes de la programmation mathématique classique peuvent s'étendre au problème présent : méthodes de gradient projeté, de directions admissibles, etc. Faisons une remarque sur les méthodes par dualité. On peut introduire le Lagrangien du problème contraint par (6.10) :

$$\mathcal{L}(u, t_1, p) = K(x(t_1), t_1) + \sum_{i=1}^r p_i \gamma_i(x(t_1), t_1) + \int_{t_0}^{t_1} L(x, u, t) dt.$$

Le calcul des dérivées de  $\mathcal{L}$  en  $u(\cdot)$  et  $t_1$  se fait exactement comme pour  $J$ , en remplaçant  $K$  par  $K + p' \gamma$  dans les équations (6.12) et (6.13).

Toutefois, comme on minimise généralement dans un ensemble fermé borné (parce que  $U$  est fermé borné), les conditions nécessaires d'optimalité ne s'écriront pas  $\partial \mathcal{L} / \partial u = 0$ . On aura  $\partial \mathcal{L} / \partial t_1 = 0$  parce que  $t_1$  est, en général, non contraint. On peut utiliser des inéquations variationnelles. Mais un résultat plus fort est donné par le principe du maximum de Pontryaguine ci-dessous.

**THÉORÈME 6.2.** Si la commande  $\bar{u}(\cdot)$  et l'instant final  $\bar{t}_1$ , sont optimaux pour le problème (6.1), (6.3), (6.10), il existe un réel positif ou nul  $\lambda_0$ , une fonction vectorielle  $\lambda(t)$  et un vecteur  $p$  satisfaisant les équations adjointes suivantes, où

$$\begin{aligned} H(x, \lambda, u, t) &= \lambda_0 L(x, u, t) + \lambda' f(x, u, t) : \\ \dot{\lambda}_k(t) &= -\frac{\partial H}{\partial x_k}(t) \\ \lambda_k(\bar{t}_1) &= \frac{\partial}{\partial x_k} (\lambda_0 K + p' \gamma)(x(\bar{t}_1), \bar{t}_1), \\ -H(\bar{t}_1) &= \frac{\partial}{\partial t_1} (\lambda_0 K + p' \gamma)(x(\bar{t}_1), \bar{t}_1), \end{aligned}$$

tels que à tous les instants  $t$  où  $\bar{u}$  est continue,

$$\forall u \in U, \quad H(x, \lambda, u, t) \geq H(x, \lambda, \bar{u}, t)$$

C'est à dire que la commande optimale fournit aussi un minimum absolu sur  $U$  de l'hamiltonien. Dans les problèmes "normaux", on peut prendre  $\lambda_0 = 1$  retrouvant ainsi les équations antérieures.

Une conséquence du principe du maximum est qu'on peut ramener la recherche d'une solution au problème de commande optimale à la solution d'un problème aux limites, obtenu en remplaçant  $H$  par

$$\bar{H}(x, \lambda, t) = \min_{u \in U} H(x, \lambda, u, t)$$

dans les equations adjointes ci-dessus. On a des conditions en  $t_1$  telles qu'évoquées ci-dessus, et en  $t_0$  :  $x(t_0) = x_0$ .

De nombreux algorithmes ont été proposés pour attaquer directement ce problème aux limites. Notamment les méthodes "de tir" où on essaye d'ajuster les conditions manquantes à un bout pour que la solution des équations différentielles correspondantes satisfasse les conditions à l'autre bout. Les méthodes de "quasilinearisation" font des linéarisations successives de l'ensemble de ces équations, pour les traiter par une méthode de type Newton.

#### 6.4) Problèmes à plusieurs agents, jeux différentiels.

On peut imaginer de nombreuses extensions du problème du paragraphe 6.1, notamment a des situations où plusieurs agents interviennent. Le système est influencé par tous :

$$\dot{x} = f(x, u_1, u_2, \dots, u_N),$$

mais leurs objectifs peuvent être plus ou moins divergents, et ils peuvent ne pas disposer des mêmes informations en temps réel. S'ils veulent minimiser le même critère on dit qu'on a un problème de théorie des équipes. S'ils veulent minimiser chacun son propre critère, on a un jeu à somme non nulle.

Un cas particulier qui a été très étudié est le cas où il y a deux "joueurs", dont nous noterons les commandes  $u$  et  $v$  :

$$\dot{x} = f(x, u, v), \quad u(\cdot) \in \Omega_u, \quad v(\cdot) \in \Omega_v, \quad (6.14)$$

et où l'un,  $u$ , cherche à minimiser le critère que son adversaire,  $v$ , veut maximiser :

$$\min_u \max_v J(u, v) = K(x(t_1), t_1) + \int_{t_0}^{t_1} L(x, u, v) dt. \quad (6.15)$$

Le concept d'optimalité retenu sera le point selle, quand il existe. On dit que  $u^*, v^*$  forment un point selle si

$$\forall u \in \Omega_u, \quad \forall v \in \Omega_v, \quad J(u^*, v) \leq J(u^*, v^*) \leq J(u, v^*).$$

Un cas typique est fourni par les problèmes de poursuite évacion. Chaque joueur pilote son propre système :

$$\dot{y} = g(y, u), \quad \dot{z} = h(z, v),$$

et l'état est  $x = (y, z)$ . La cible peut être de la forme

$$C = \left\{ \begin{pmatrix} y \\ z \end{pmatrix} \mid |y - z| \leq \ell \right\}$$

par exemple (les deux joueurs sont à une distance inférieure ou égale à  $\ell$ ) et le critère le "temps de capture". On appelle le minimiseur "poursuivant" et le maximiseur "fugitif".

Les problèmes de jeux différentiels se révèlent beaucoup plus difficiles que ceux de commande optimale. Dans les jeux de poursuite évacion, par exemple, avant de chercher à minimiser le temps de capture, le poursuivant doit d'abord savoir s'il peut s'assurer cette capture. Suivant les problèmes, pour certains états initiaux le fugitif pourra au contraire s'assurer de n'être jamais capturé. Il y a donc d'abord un jeu "qualitatif" : quels sont, pour chaque joueur, les états initiaux d'où il peut s'assurer de "gagner", et par quelle stratégie. Puis, pour les états "capturés", le jeu "quantitatif" posé au départ. Les zones de victoire de chacun des

joueurs sont séparées, dans l'espace d'états, par des "barrières", surface qu'aucun des deux joueurs ne peut forcer l'état à traverser dans le sens qui lui serait favorable (vers sa propre zone de victoire).

Le principal outil disponible est ici le parallèle à deux joueurs de l'équation de Hamilton Jacobi Carathéodory, connu sous le nom d'équation d'Isaacs. Les méthodes variationnelles ne s'adaptent pas. En effet, dans tout ce discours, il est sous entendu que les deux joueurs ont la possibilité d'utiliser des stratégies en boucle fermée :

$$u(t) = \varphi(x(t), t), \quad v(t) = \psi(x(t), t). \quad (6.16)$$

De même que le chat ne poursuit pas la souris les yeux fermés en se fondant sur l'endroit où elle "devrait" aller, de même l'information disponible aux joueurs joue un rôle fondamental dans toute cette théorie. Les choses deviennent encore infiniment plus compliquées si les joueurs ne disposent que d'une information imparfaite. Au point que très peu de résultats sont disponibles pour de tels problèmes (un résultat a été obtenu en 1990 permettant notamment de traiter la commande  $H_\infty$  en feedback de sortie. Voir paragraphe 5.6).

Comme on peut s'y attendre, un cas assez simple est fourni par le problème linéaire quadratique, où les équations (6.14) et (6.15) prennent la forme suivante:

$$\dot{x} = Ax + Bu + Ev, \quad (6.17)$$

$$J = x'(t_1)Q_1x(t_1) + \int_{t_0}^{t_1} (x'Q(t)x + u'R(t)u - v'S(t)v) dt. \quad (6.18)$$

avec  $t_1$  donné et sans contrainte sur  $x(t_1)$ . La solution fait intervenir une équation de Riccati sur une matrice symétrique  $n \times n$ ,  $P(t)$  (équation dont un cas particulier intervient au paragraphe 5.6).

$$\dot{P} + AP + PA' - P(BR^{-1}B' - ES^{-1}E')P + Q = 0, \quad P(t_1) = Q_1. \quad (6.19)$$

On a le résultat suivant :

**THÉORÈME 6.3.** Le jeu (6.17)(6.18) a un point selle avec des stratégies de la forme (6.16) à coefficients bornés si et seulement si l'équation de Riccati (6.19) a une solution sur  $[t_0, t_1]$ . Dans ce cas, les stratégies optimales sont données par

$$u(t) = -R^{-1}(t)B'(t)P(t)x(t), \quad v(t) = S^{-1}(t)E'(t)P(t)x(t)$$

et le coût optimal par  $x_0'P(t_0)x_0$ .

Mais là encore la théorie est plus complexe qu'en commande optimale : il peut se faire que l'équation de Riccati diverge pour  $t \rightarrow t_1^* \in (t_0, t_1)$  mais que le jeu ait encore un point selle si on autorise un feedback dont le gain diverge aussi au voisinage de  $t_1^*$ , ce feedback assurant que  $x(t)$  reste tel que la commande correspondante reste bornée, et en tous cas de carré sommable. Il faut bien sûr un autre outil que l'équation de Riccati pour trouver les stratégies entre  $t_0$  et  $t_1^*$ .

## 7) Aperçu sur les systèmes non linéaires

### 7.1) Introduction

Le présent article fait une très large part aux systèmes linéaires, conformément à ce qui était annoncé dans l'introduction. En fait, il n'a été question, jusqu'à présent, que de ces derniers mis à part dans le paragraphe consacré à la commande optimale où l'hypothèse de linéarité n'est pas essentielle pour développer la théorie. Encore convient-il d'observer que les lois de commande obtenues alors ne sont en général pas, dans le cas non-linéaire, obtenues sous forme de feedback. Cependant, le modèle linéaire présente des limitations évidentes, et ne saurait suffire à décrire globalement la plupart des systèmes. En d'autres termes, chaque fois que le point de fonctionnement d'un système ne reste pas au voisinage d'un régime nominal (la notion de voisinage dépend bien sûr du système lui-même) ou que le linéarisé du système est singulier, ce linéarisé ne permet plus d'en décrire convenablement le comportement et on a affaire à un problème non-linéaire. Cela se produit, par exemple, lorsqu'on veut amener le système d'un état  $x$  à un état  $y$  et que le trajet  $xy$  n'est pas suffisamment petit. Une méthode souvent envisagée pour obvier cela est de déterminer une loi de commande

nominale  $u(t)$  pour relier  $x$  à  $y$  par les techniques du contrôle optimal décrites au paragraphe précédent, puis de linéariser le système gouverné par cette commande au voisinage de la trajectoire obtenue pour retrouver un problème de commande linéaire non stationnaire qu'on peut à son tour discrétiser ou encore traiter par des techniques de "gain scheduling". Pour ingénieuse qu'elle soit, il serait vain de penser que cette technique dispense de faire face à de vraies questions non-linéaires. Tout d'abord, le problème de contrôle optimal en question ne peut avoir de solution que si  $y$  est accessible à partir de  $x$ , ce qui pose le problème de la contrôlabilité du système. Par ailleurs, des singularités le long de la trajectoire envisagée peuvent, et souvent doivent, se produire (une singularité est ici un point où le linéarisé serait non commandable ou non observable). Enfin, le calcul et la mémorisation de la trajectoire  $u$  discrétisée ainsi que la linéarisation en des points suffisamment denses sur la trajectoire peuvent être tout simplement hors de portée au plan calculatoire, notamment lors d'applications en temps réel. Nous sommes ainsi conduits à examiner plus avant les analogues non-linéaires des notions systémiques introduites auparavant. Et immédiatement, nous nous heurtons à une première difficulté qui est que *non-linéaire* ne constitue guère une définition. Il semble en fait que la détermination d'une ou plusieurs classes de systèmes à la fois suffisamment générales pour couvrir substantiellement les applications et suffisamment simples pour se prêter globalement à une étude quantitative approfondie soit une question fondamentale de la théorie des systèmes pour laquelle il n'existe pas, à l'heure actuelle, de véritable réponse. Du moins la recherche d'un paradigme non-linéaire a-t-elle conduit, chemin faisant, au développement de plusieurs pans de théorie à la fois beaux et profonds dont il n'est pas possible de tous rendre compte ici. Nous nous contenterons de commenter brièvement certains concepts ouvrant des perspectives depuis la théorie linéaire.

Une des généralisations les plus naturelles des systèmes linéaires, et historiquement l'une des premières est le concept de *série de Volterra*. Il s'agit d'une application entrée-sortie  $u \rightarrow y$  du type

$$y(t) = \sum_{i=1}^{\infty} \int_{-\infty}^{+\infty} h_i(\tau_1, \dots, \tau_i) u(t - \tau_1) \dots u(t - \tau_i) d\tau_1 \dots d\tau_i,$$

la causalité se traduisant par le nullité de  $h_i$  lorsqu'un de ses arguments est négatif. Lorsque la série ci-dessus est finie, on parle de *système polynomial*. Ce développement en série d'un système est réminiscent du développement de Taylor d'une fonction, le premier terme étant un système linéaire. Les problèmes de convergence d'une telle série, cependant sont délicats. Un des attraits de la représentation de Volterra est qu'elle se prête à une généralisation de la notion de fonction de transfert en utilisant la transformée de Laplace multivariable, mais cette construction n'est pas aussi efficace que son analogue linéaire. En fait, il semble que la généralisation de la notion de fonction de transfert soit plutôt à chercher dans l'approche algèbre-différentielle que nous aurons l'occasion de mentionner plus loin, et nous ne discuterons pas ici davantage de descriptions entrée-sortie. Ce choix nous amènera à passer sous silence un certain nombre d'autres développements théoriques qui ont eu une influence assez profonde sur la discipline, en particulier le concept de série génératrice mieux adapté que les séries de Volterra à l'étude des problèmes structurels, et qui établit des liens profonds entre la théorie des systèmes et celles des automates et des groupes formels.

Un modèle déjà très général, beaucoup trop général à vrai dire pour théoriser une synthèse observateur-correcteur, est fourni par (2.1). On peut au besoin rendre un tel système autonome en adjoignant à  $x$  la variable  $t$  dont la dynamique est  $\dot{t} = 1$  de sorte que, quittes à redéfinir  $f$ ,  $h$  et  $x$ , on a une équation d'état du type

$$\dot{x} = f(x, u), \quad y = h(x, u), \quad (7.1)$$

$$x(t) \in \mathbb{R}^n, \quad u(t) \in \mathbb{R}^m, \quad y(t) \in \mathbb{R}^p,$$

où la dépendance explicite en  $t$  a disparue. Les fonctions  $f$  et  $h$  sont supposées lisses (i.e. indéfiniment dérivables par rapport à leurs arguments) et, en maintes occasions, il est même important qu'elles soient analytiques (réelles), auquel cas on dit que le système lui-même est analytique. Il est enfin de bonnes raisons de limiter la classe des entrées  $u(t)$  admissibles, non seulement pour assurer l'existence d'une solution à (7.1) c'est à dire la validité du modèle lui-même, mais aussi pour éviter certaines pathologies. En effet, si nous prenons comme exemple le système

$$\dot{x} = xu, \quad u(t) \in \mathbb{R} \quad x(t) \in \mathbb{R} \quad (7.2)$$

initialisé en  $x(0) = 0$ , il est clair que la commande  $u(t) = 2/t$  amène à  $x(1) = 1$  à l'instant 1, mais aucune commande sommable ne permet de quitter l'état zéro. Pour éviter ce genre de problèmes, on suppose que les commandes  $u(t)$  sont localement bornées (et bien sûr mesurables) comme fonctions de  $t$ .

En dépit de sa généralité, la formulation (7.1) comporte déjà trois restrictions notables. Tout d'abord, on a supposé que le contrôle  $u(t)$  n'était pas borné, ce qui est physiquement peu réaliste. Pour faire percevoir à quel point une contrainte du type  $\|u(t)\| \leq M$  peut changer la situation, observons seulement que les seuls systèmes linéaires complètement commandables avec cette restriction sont ceux dont les pôles sont imaginaires purs. Deuxièmement, nous nous sommes limités au cas où  $x$  peut varier dans  $\mathbb{R}^n$ , alors qu'il est souvent nécessaire de le faire évoluer sur des variétés plus générales, par exemple dans les problèmes mécaniques où le repère lié au corps évolue dans le groupe des rotations. A cet égard, (7.1) peut être interprété comme une équation d'évolution locale, c'est à dire tant que  $x(t)$  reste dans le domaine d'une carte, mais certaines caractéristiques de la variété sur laquelle évolue l'état peuvent être elles-mêmes pertinentes pour le contrôle, en particulier la compacité ou une structure de groupe de Lie. Nous n'en dirons cependant rien. Troisièmement, et dans le même ordre d'idées, on a postulé que  $f$  est définie et régulière sur tout  $\mathbb{R}^n$  mais cela n'est pas très souvent le cas, car il existe fréquemment des "points singuliers" où l'analyse est délicate. Cette présence de singularités, c'est à dire de points où le système ne peut être décrit localement par une équation du type (7.1) où toutes les quantités sont lisses, est d'ailleurs une caractéristique majeure des problèmes non linéaires, et un des principaux obstacles à l'obtention de résultats globaux.

Après ce préambule, nous consacrerons le prochain paragraphe à une approche assez générale de la notion de contrôlabilité pour le système (7.1) fondée sur des notions classiques de géométrie différentielle. Il s'agit historiquement du premier abord systématique de la question.

### 7.1) Contrôlabilité, observabilité, minimalité

Introduisons tout d'abord quelques notions. Un champ de vecteur, sur  $\mathbb{R}^n$ , sera une application  $X : \mathbb{R}^n \rightarrow \mathbb{R}^n$  associant à tout point  $x$  un vecteur  $X(x)$  et ce de façon lisse. L'équation différentielle

$$\dot{x} = X(x) \quad x(0) = x_0$$

admet une unique solution dont on supposera pour simplifier, mais ce ne serait indispensable que pour l'unicité d'une réalisation minimale, qu'elle est définie à tout instant (i.e. tous les champs qui interviendront seront supposés complets) et dont la valeur à l'instant  $t$  sera notée  $\exp(tX)x_0$ . Si  $X$  est un champ de vecteurs de composantes  $(c_1(x), c_2(x), \dots, c_n(x))$ , sa dérivée est définie en chaque point  $x = (x_1, \dots, x_n)$  comme étant la matrice carrée  $DX(x)$  dont la composante  $i, j$  est  $\partial c_i / \partial x_j$ . Si, à présent,  $X$  et  $Y$  sont deux champs de vecteurs, leur *crochet de Lie* est un troisième champ de vecteurs dont la valeur au point  $x$  est

$$[X, Y] = DX(x).Y(x) - DY(x).X(x),$$

où le point désigne la multiplication des matrices. Pour illustrer cette notion, supposons que, dans le système (7.1) se trouvant en l'état  $x_0$ , on utilise une commande constante égale à  $u$  pendant le laps de temps  $t$  puis une commande constante  $v$  pendant le même laps de temps  $t$ . Le système se trouve ainsi amené en un état  $x_1(t)$ . Si on appelle  $X$  le champ de vecteurs  $f(x, u)$  et  $Y$  le champ de vecteurs  $f(x, v)$ , on a

$$x_1(t) = \exp(Yt)\exp(Xt)x_0.$$

De façon symétrique, si on applique d'abord la commande  $v$  puis la commande  $u$ , le système se trouve amené en l'état

$$x_2(t) = \exp(Xt)\exp(Yt)x_0.$$

La différence  $x_1 - x_2$  est une fonction lisse de  $t$  dont un calcul classique montre que, en  $t = 0$ , sa dérivée est nulle cependant que sa dérivée seconde est  $2[X, Y]$ . Cet exemple permet de percevoir en quoi le crochet de Lie est relié à l'effet du séquençement de commandes constantes par morceaux. L'hypothèse que  $u(t)$  est localement borné entraînant qu'il peut être approché, en un certain sens, par des contrôles constants par morceaux, il s'ensuit des relations générales entre crochet et contrôle. Pour énoncer un résultat plus précis il nous faut introduire à présent quelques définitions supplémentaires. L'ensemble des champs de vecteurs est manifestement un espace vectoriel sur  $\mathbb{R}$ , et un sous-espace vectoriel qui est stable par formation du crochet

de Lie est appelé algèbre de Lie. Lorsqu'on évalue en un point  $x$  tous les champs  $X$  d'une algèbre, on obtient une collection  $X(x)$  de *vecteurs de  $\mathbb{R}^n$*  qui forment un espace vectoriel dont la dimension est appelée *rang* de l'algèbre au point  $x$ . Un idéal  $\mathcal{I}$  d'une algèbre de Lie  $\mathcal{L}$  est une sous-algèbre absorbante pour le crochet, c'est à dire telle que si  $X \in \mathcal{I}$  et  $Y \in \mathcal{L}$ , alors  $[X, Y] \in \mathcal{I}$ . Par exemple, l'espace engendré par tous les crochets  $[X, Y]$  lorsque  $X$  et  $Y$  décrivent  $\mathcal{L}$  est manifestement un idéal appelé algèbre dérivée de  $\mathcal{L}$ . Etant donnée une collection  $\mathcal{X}$  de champs de vecteurs, on appelle algèbre de Lie engendrée par  $\mathcal{X}$  la plus petite sous algèbre de Lie contenant  $\mathcal{X}$ , qui n'est autre que l'espace vectoriel engendré par les éléments de  $\mathcal{X}$  et tous les crochets itérés que l'on peut effectuer à partir de ceux-ci. On appellera *algèbre de commandabilité* du système (7.1) et on notera  $\mathcal{L}(f)$  l'algèbre de Lie engendrée par la collection de champs  $f(x, u)$  lorsque  $u$  décrit  $\mathbb{R}^m$ . On définira de plus  $\mathcal{L}_0(f)$  comme étant la sous-algèbre de  $\mathcal{L}(f)$  formée de toutes les sommes

$$\sum_{k=1}^N \lambda_k f(x, u_k) + Y,$$

où  $u_1, \dots, u_N$  est une suite finie arbitraire d'éléments de  $\mathbb{R}^m$ , les réels  $\lambda_k$  étant astreints à être de somme nulle, cependant que  $Y$  décrit l'algèbre dérivée de  $\mathcal{L}(f)$ . Il est clair que  $\mathcal{L}_0(f)$  est un idéal qu'on nommera *idéal de commandabilité* du système. La nécessité d'introduire à la fois  $\mathcal{L}(f)$  et  $\mathcal{L}_0(f)$  provient de la distinction, occultée dans le cas des systèmes linéaires stationnaires, qu'il faut faire entre la possibilité d'atteindre un état et celle de l'atteindre en un temps déterminé. De manière plus précise, on dira que l'état  $y$  est accessible en temps  $T > 0$  à partir de l'état  $x$  s'il existe une commande  $u(t)$  engendrant une trajectoire  $x(t)$  de (7.1) telle que  $x(0) = x$  et  $x(T) = y$ . On dira simplement que  $y$  est accessible à partir de  $x$  s'il existe un  $T > 0$  tel que  $y$  soit accessible en temps  $T$ . On notera  $\mathbf{A}(x)$  l'ensemble des états accessibles à partir de  $x$  et  $A(x, T)$  ceux qui le sont en temps  $T$ .

**THÉOREME 7.1.** Etant donné le système (7.1), une condition suffisante pour que  $\mathbf{A}(x)$  soit d'intérieur non vide est que le rang de  $\mathcal{L}(f)$  au point  $x$  soit égal à  $n$ , et une condition suffisante pour que  $A(x, T)$  soit d'intérieur non vide pour un certain (et alors pour tout)  $T > 0$  est que le rang de  $\mathcal{L}_0(f)$  au point  $x$  soit  $n$ . Ces conditions sont également nécessaires si le système est analytique.

L'énoncé ci-dessus illustre bien l'importance de l'analyticité qui rend nécessaire une condition d'ores et déjà suffisante parce que la rigidité analytique fait que toute l'information est donnée par la suite des dérivées à l'origine laquelle est, *grosso modo*, contenue dans l'algèbre de Lie au point initial. Mentionnons aussi que ce résultat est intimement lié au classique théorème d'intégrabilité de Frobenius de la géométrie différentielle. Sa signification principale est bien sûr que, sous certaines conditions, l'ensemble des états accessibles à partir d'un état initial donné a "de l'épaisseur". Par exemple, on vérifie aisément que l'algèbre de commandabilité du système (7.2) au point 0 est nulle, en accord avec le fait que l'on ne peut quitter ce point. Lorsque  $\mathbf{A}(x)$  (resp.  $A(x, T)$ ) sont d'intérieur non vide pour tout  $x$ , on dit que le système est complètement accessible (resp. fortement complètement accessible). Ainsi, le théorème 7.1 propose le critère d'accessibilité suivant:

*Un système analytique est complètement accessible (resp. fortement complètement accessible) si et seulement si son algèbre (resp. idéal) de commandabilité est de plein rang en tout point.*

On peut montrer que la forte accessibilité est liée à la complète commandabilité générique du système linéarisé le long de la trajectoire. Ceci ne suffit pas, cependant, à trancher de la propriété de *commandabilité*, qui se définit comme la possibilité d'aller de tout état  $x$  à tout état  $y$ . Ce défaut est malheureusement dans la nature des choses, et traduit l'impossibilité pour un critère algébrique d'accessibilité à prendre en compte *l'irréversibilité du temps*. Un exemple trivial mais instructif est donné par le système

$$\dot{x} = u^2, \quad u(t) \in \mathbb{R} \quad x(t) \in \mathbb{R},$$

dont il est clair qu'il est fortement accessible, mais non commandable puisque  $x$  ne peut décroître. Notons que, dans ce cas,

$$A(x_0, T) = \{x \in \mathbb{R}; x \geq x_0\},$$

de sorte que  $x_0$  n'appartient pas à l'intérieur des états qui lui sont accessibles. En fait, la relation d'accessibilité n'est pas ici *symétrique*, autrement dit le fait que  $y$  soit accessible à partir de  $x$ , n'entraîne pas que  $x$  soit accessible à partir de  $y$ . Cette symétrie de la relation d'accessibilité qui, d'une certaine manière, traduit la

capacité du système à “remonter le temps”, est la notion globale essentielle en commandabilité comme le montre le résultat suivant:

**THÉOREME 7.2.** Le système analytique (7.1), est commandable si et seulement si il est fortement accessible et si en outre la relation d’accessibilité est symétrique.

Pour satisfaisant qu’il paraisse, le théorème 7.2 est en réalité assez tautologique, car aucune méthode un tant soit peu générale ne permet de décider de la dernière propriété qui, si elle a lieu, provient usuellement de certaines symétries de la fonction  $f(x, u)$  qu’il convient d’analyser dans chaque cas. Cette situation illustre bien la difficulté qu’il ya a à obtenir des énoncés à la fois globaux et puissants dans la théorie non linéaire. A l’inverse, il existe des résultats (plus ou moins effectifs) unifiant la plupart des propriétés connues de commandabilité *locale* mais, plutôt que de les discuter, nous nous contenterons pour conclure d’examiner quelques exemples.

Si le système est linéaire, c’est à dire si  $f(x, u) = Ax + Bu$ , il est évidemment analytique et il est facile de voir que l’idéal de commandabilité engendre en chaque point un espace vectoriel qui est l’image de la matrice  $(B, AB, A^2B, \dots)$ . Celle-ci doit donc être de plein rang pour que le système soit commandable. Mais si c’est le cas,  $\mathbf{A}(0)$  qui est manifestement un espace vectoriel doit être d’intérieur non-vidé et donc égal à  $\mathbb{R}^n$ . Tout point est donc accessible à partir de 0 et il est assez facile de voir qu’alors 0 est aussi accessible à partir de tout point, ce qui permet de retrouver le critère de commandabilité du théorème 3.6.

Si le système est *linéaire en la commande*, c’est à dire si

$$f(x, u) = \sum_{k=1}^m u_k f_k(x),$$

où les  $u_k$  sont les coordonnées de  $u$  et les  $f_k$  des champs de vecteurs sur  $\mathbb{R}^n$ , il est facile de voir que l’accessibilité est symétrique et que  $\mathcal{L}(f) = \mathcal{L}_0(f)$  n’est autre que l’algèbre de Lie engendrée par les  $f_k$ . La condition que celle-ci soit de rang plein en tout point est suffisante pour entraîner la commandabilité, et par ailleurs nécessaire si les  $f_k$  sont analytiques.

Une classe nettement plus générale, couvrant beaucoup d’applications, est celle des systèmes *affines en la commande* pour lesquels

$$f(x, u) = f_0(x) + \sum_{k=1}^m u_k f_k(x), \tag{7.3}$$

le terme  $f_0$  représentant l’inertie du système. Notons que lorsque l’algèbre de Lie engendrée par les  $f_k$  pour  $k \geq 1$  est de rang plein en tout point, le système est commandable parce que l’on peut dominer l’effet de dérive de  $f_0$  en utilisant le contrôle, complet d’après l’exemple précédent, que l’on possède grâce aux  $f_k$  pour  $k \geq 1$ . Ceci fait évidemment usage du caractère non borné de  $u$  et n’est donc pas très réaliste, mais fournit une condition de commandabilité bien sûr plus exigeante que la condition de forte accessibilité qui dit ici que l’idéal engendré par les  $f_k$  pour  $k \geq 1$  dans l’algèbre de Lie engendrée par tous les  $f_j$  est de rang plein en tout point. Cette condition de forte accessibilité n’est, quant à elle, pas suffisante en général pour assurer la commandabilité mais peut le devenir sous des instances particulières et notamment lorsque  $f$  possède certaines imparités ou encore lorsque les solutions de l’équation différentielles  $\dot{x} = f_0(x)$  sont presque périodiques, ce qui signifie que la dynamique propre du système revient approximativement sur elle-même au bout d’un certain temps. Dans ce dernier cas, on peut même supposer les contrôles bornés. Bien que nous sortions ici du cadre formel de notre exposé, observons que cette circonstance est assez susceptible de se produire si l’état  $x$ , au lieu d’évoluer dans  $\mathbb{R}^n$ , appartient à une variété compacte. Historiquement, ce fut un des premiers succès de la théorie que d’analyser ainsi la gouvernabilité d’un satellite en rotation, problème qui échappait aux techniques de linéarisation.

Il est historiquement également quelque raison de réserver une mention spéciale à certains systèmes affines particuliers qui ont été beaucoup étudiés et qu’on appelle *systèmes bilinéaires*. Ce sont ceux pour lesquels les champs  $f_k$  sont des champs affines du type  $A_k x + b_k$  où les  $A_i$  sont des matrices constantes de taille convenable et les  $b_i$  des vecteurs de  $\mathbb{R}^n$ , cependant que l’équation de sortie est linéaire:  $y = Cx$ . Pour ces systèmes, les calculs de crochets nécessaires pour tester l’accessibilité se ramènent à ceux de commutateurs ordinaires de matrices. La classe bilinéaire possède une propriété remarquable, à savoir que tout système de type (7.1), lorsqu’on considère son comportement à partir d’un état initial donné sur une intervalle de temps

fini fixé, peut être approché arbitrairement bien par un système bilinéaire au sens de la convergence uniforme des sorties. Si on regarde les modèles linéaires comme des développements au premier ordre en  $x$  et  $u$  de (7.1) et les modèles bilinéaire comme des développements d'ordre deux, ce résultat est assez surprenant en ce sens que l'ordre deux suffit à obtenir une approximation dynamique aussi fine que l'on veut en un certain sens. Cette propriété de densité, jointe à certaines autres propriétés comme la rationalité de leur fonction de transfert, a fait des systèmes bilinéaires une classe prototypique mais, et ceci est peut-être dû à cela, ils ne sont pas notoirement plus aisés à manipuler. Leur contrôle, par exemple, n'est pour l'instant relativement élucidé que si  $n = 2$ .

La notion d'*observabilité* pour le système (7.1), que nous étudions à présent, n'est pas la complète duale de la commandabilité que l'on pourrait attendre. Plus directe à manier sous certains aspects comme la minimalité, elle est plus délicate sous d'autres. La première raison en est que l'ensemble des états *indistinguables* d'un état  $x_0$  donné, qui est par définition l'ensemble des états  $x$  tels que l'application d'une commande  $u$  à partir de  $x_0$  et à partir de  $x$  donne lieu à la même sortie et ce quelle que soit la commande choisie, n'est ni connexe, ni de dimension constante en général, même dans le cas analytique. Pour en décomposer localement la structure, on introduit la notion, par ailleurs assez artificielle, de forte indistinguabilité selon laquelle deux états en relation doivent pouvoir être reliés par une courbe continues d'états indistinguables. Pour les systèmes linéaires, il n'y a bien sûr pas de différence avec la notion habituelle puisqu'alors deux états sont indistinguables si et seulement si leur différence appartient au noyau de la matrice d'observabilité. Mais pour le système

$$\dot{x} = u \quad y = \sin x \quad u(t) \in \mathbb{R} \quad x(t) \in \mathbb{R}, \quad (7.4)$$

$x$  et  $x + 2\pi$  sont indistinguables sans l'être fortement. L'ensemble des états fortement indistinguables d'un état  $x_0$  est, *grosso modo* la plus grande hypersurface connexe de l'espace d'état qui contient  $x$ , qui est invariante par la dynamique du système, et qui est tangente au noyau de la dérivée  $\partial h / \partial x$  de  $h$ . Ainsi, si on évolue sur cette surface, les valeurs prises par  $h$  ne sont influencées que par  $u$ , et non par  $x$ , et en particulier ne dépendent pas du point initial. Le fait que sa définition soit différentielle va avoir pour conséquence que sa dimension peut être évaluée avec une condition de rang duale de l'accessibilité. Pour donner formellement corps à cette notion introduisons une collection de fonctions en définissant d'abord  $h_k(x, u)$  comme les composantes de  $h$  pour  $1 \leq k \leq p$ , puis l'ensemble  $\mathcal{F}_0$  des fonctions de  $x$  du type  $h_k(\cdot, u)$  que l'on obtient en donnant à  $u$  toutes les valeurs réelles. En considérant les dérivées des éléments de  $\mathcal{F}_0$  dans les directions  $f(x, v)$  où  $v$  parcourt  $\mathbb{R}$ , c'est à dire les fonctions du type

$$\sum_{j=1}^n \frac{\partial g}{\partial x_j} f_j(x, v), \quad g \in \mathcal{F}_0$$

où  $f_j$  indique la  $j$ ème composante de  $f$ , on fabrique une nouvelle collection  $\mathcal{F}_1$  de fonctions. On itère ce processus pour obtenir une suite  $\mathcal{F}_i$  de collection de fonctions, et on pose finalement  $\mathcal{F} = \cup \mathcal{F}_i$ . L'ensemble des différentielles de toutes les fonctions de  $\mathcal{F}$  engendre en chaque point  $x$  un sous espace du dual de  $\mathbb{R}^n$ , dont la dimension sera appelée *rang* de  $\mathcal{F}$  au point  $x$  (la dualité voudrait ici que l'on introduise l'idéal différentiel engendré par ces formes et qu'on parle d'*idéal d'observabilité*, mais il serait un peu lourd d'introduire cette notion). On a alors le

**THÉOREME 7.3.** Etant donné le système (7.1), une condition suffisante pour que deux états distincts soient toujours fortement distinguables est que le rang de  $\mathcal{F}$  soit égal à  $n$  en tout point  $x$ . Cette condition est également nécessaire si le système est analytique.

Dans le cas linéaire, ceci redonne le critère d'observabilité (3.17). Si nous convenons de dire que le système est *localement observable* (c'est effectivement à cela que cela revient dans le cas analytique) lorsque la condition de rang du théorème est satisfaite, l'analogue du théorème 7.2 serait *le système analytique (7.1) est observable si et seulement si il est localement observable et si deux états indistinguables peuvent être joints par une courbes d'états indistinguables*, un énoncé à nouveau plutôt tautologique.

Nous sommes à présent en mesure d'aborder la question de la minimalité, *que nous n'aborderons que dans le cas analytique*. Ici, la dissymétrie entre observabilité et commandabilité devient apparente. En effet, la commandabilité ne peut s'obtenir qu'en se restreignant à l'ensemble des états accessibles à partir de l'état initial, qui n'est pas un objet suffisamment régulier en général pour que l'on puisse définir dessus un système dynamique du même type que celui dont on est parti. Nous sommes donc amenés à restreindre nos exigences

en demandant simplement à un système minimal d'être accessible, ce qui astreint l'ensemble des états à se déplacer sur un objet géométrique régulier, à savoir une variété, de dimension aussi faible que possible (qu'en général l'état ne décrit pas dans son ensemble à moins que le système ne soit commandable). A l'inverse, on peut demander à la réalisation minimale d'être observable, parce que la relation d'indistinguabilité est suffisamment régulière pour que le quotient d'une variété soit encore une variété. Il faut cependant noter que, dans tout les cas, on quitte à nouveau le cadre formel de notre exposé parce que la réalisation minimale de (7.1) n'a aucune raison d'évoluer sur  $\mathbb{R}^k$ . L'espace d'état de la réalisation minimale de (7.4), par exemple, est un cercle. Une *réalisation minimale* de (7.1) sera donc un système analytique du même type mais où  $x$  évoluera sur une variété analytique  $M$ , qui sera accessible et observable, et tel qu'il existera une application lisse  $\mathbb{R}^n \rightarrow M$  qui commutera avec l'action des commandes et l'évaluation de la sortie. Deux systèmes, définis tous deux sur des variétés  $M_1$  et  $M_2$ , et tels qu'il existe un isomorphisme des espaces d'état  $M_1 \rightarrow M_2$  commutant avec l'action des commandes et l'évaluation des sorties, sont par ailleurs dits isomorphes. On peut alors énoncer le théorème de réalisation minimale

**THÉORÈME 7.4.** Tout système analytique du type (7.1), initialisé en un état  $x_0$  donné, admet une réalisation minimale qui est unique à un isomorphisme près.

En conclusion de ce paragraphe, il importe de remarquer que si on a réussi à analyser quelque peu les possibilités de commande et d'observation d'un système, rien n'a été dit sur *la manière* d'y parvenir, même si l'on a pu prouver que c'était possible. En ce qui concerne l'observation, il est une différence d'avec le cas linéaire qu'il importe de mentionner ici, à savoir que deux états distinguables peuvent ne pas être distingués par toute entrée. Cela conduit à introduire la notion *d'entrée universelle*, c'est à dire une entrée qui suffirait à elle seule à distinguer entre eux tous les états distinguables. On peut prouver qu'une telle entrée existe, mais sa construction n'est pas effective en général, ce qui est source de difficultés lorsqu'il s'agit de concevoir un observateur. Certaines constructions existent, notamment pour les systèmes bilinéaires, en dimension finie. D'autres encore linéarisent le problème en dimension infinie. De ce domaine actif de la recherche nous ne dirons rien. En ce qui concerne le contrôle, il n'existe pas à ce jour de méthode relativement générale pour définir une commande menant d'un état initial à un état donné, quoique des avancées récentes aient été effectuées dans le cas linéaire en la commande, et donc en particulier les systèmes commandés en vitesse. Ceci contribue à expliquer l'importance, dans ce contexte, du contrôle optimal traité au chapitre 6 qui reste un des moyens principaux de calculer une commande en l'exhibant comme solution d'un problème de minimisation que souvent, par ailleurs, on ne souhaitait pas particulièrement résoudre. Le critère, à cet égard, est donc le plus fréquemment choisi de manière à faciliter, dans la mesure du possible, la résolution. Les problèmes posés par la nature en boucle ouverte de ces solutions nous ramènent aux considérations premières de ce chapitre, et nous nous contenterons de décrire rapidement, lors des paragraphes suivants, certaines versions affaiblies du problème de commande général.

## 7.2) Approche géométrique, linéarisation.

Le langage de la géométrie différentielle introduit ci-dessus permet de traduire de façon quasi-systématique les énoncés de l'approche géométrique linéaire en leurs homologues non linéaires. Les énoncés ainsi obtenus n'ont cependant de valeur que localement en général. Nous nous contenterons de discuter un analogue du théorème 3.12 concernant le rejet des perturbations pour les systèmes analytiques affines en la commande. On considère un système analytique du type (7.3) avec une équation de sortie

$$\begin{aligned} \dot{x} &= f_0(x) + \sum_{k=1}^m u_k f_k(x) + \sum_{j=1}^r q_j g_j(x) \\ y &= h(x) \end{aligned} \tag{7.5}$$

où les  $q_j$  sont des perturbations inconnues affectant la dynamique par le biais des champs de vecteurs  $g_j$ . On se demande à présent s'il existe un feedback du type

$$u_k = \alpha_k(x) + \sum_{i=1}^m \beta_{k,i}(x) v_i \tag{7.6}$$

qui rende la sortie  $y$  indépendante des perturbations  $q_j$ , où les  $\alpha_k$  et  $\beta_{k,i}$  sont des fonctions lisses qu'il s'agit de choisir convenablement. Notons qu'il y a là une petite généralisation du problème présenté au

chapitre 3 pour lequel les  $\beta_{k,i}$  étaient nuls alors qu'on cherche ici à conserver si possible des variables de commande pour le système bouclé qui sont précisément les  $v_i$ . On dit que le rejet des perturbations est réalisé *localement* en  $x_0$  si on a pu définir un tel feedback lorsque  $x$  reste dans un voisinage de  $x_0$ . Il faut noter que cette formulation suppose que l'on a accès à  $x$ , donc qu'on a pu construire un observateur. Afin de donner l'énoncé correspondant à 3.12, il nous faut décrire la correspondance entre les objets de l'approche géométrique linéaire et leurs homologues non linéaires. Ce dictionnaire est obtenu en remplaçant les espaces vectoriels par des *distributions* de vecteurs –il faut prendre garde de ce qu'il ne s'agit pas des mêmes objets que les distributions de l'analyse. Une distribution  $D$  est ici une application qui associe à chaque  $x$  un sous-espace vectoriel  $D(x)$  de  $\mathbb{R}^n$ . Toute distribution sera tacitement supposée analytique, c'est à dire telle qu'il existe une famille de champs de vecteurs analytiques  $(b_s)_{s \in S}$  de sorte que, pour tout  $x$ , l'espace vectoriel engendré par les  $b_s(x)$  soit égal à  $D(x)$ . On dit par abus de langage qu'un champ de vecteurs  $b$  appartient à  $D$  si  $b(x) \in D(x)$  pour tout  $x$ , et la distribution sera dite *involutive* si le crochet de deux champs qui y appartiennent y appartient encore. La dimension de  $D(x)$  en tant qu'espace vectoriel sur  $\mathbb{R}$  est appelée *le rang* de  $D$  au point  $x$ . Si  $b$  est un champ de vecteurs, on note  $ad_b$  l'application qui à un champ de vecteurs  $X$  associe le champ  $[b, X]$ . La même terminologie est adoptée sans distinction lorsque tous ces objets sont définis sur un ouvert de  $\mathbb{R}^n$  plutôt que  $\mathbb{R}^n$  en entier. La notion d'espace vectoriel invariant par une application linéaire est remplacée par celle de distribution invariante par un champ de vecteurs: on dit que  $D$  est *b invariante* si  $ad_b(D) \subset D$ . Il est facile de voir que si  $D$  est la distribution constante qui associe à tout point  $x$  un espace vectoriel  $\mathcal{V}$  et si  $b$  est le champ  $x \rightarrow Ax$  où  $A$  est une application linéaire, on retrouve la notion d'espace  $A$ -invariant. Si on nomme  $\mathcal{X}$  l'ensemble des champs de vecteurs  $X$  qui annulent  $h$ , c'est à dire tels que

$$\sum_{i=1}^n \frac{\partial h}{\partial x_i} X_i$$

soit la fonction identiquement nulle,  $X_i$  désignant la  $i$ -ème composante de  $X$ , l'espace vectoriel engendré en chaque point par les champs de  $\mathcal{X}$  évalués en ce point est une distribution involutive que nous noterons  $\text{Ker}h$ . Substituons à présent (7.6) dans (7.5). Le système bouclé s'écrit:

$$\begin{aligned} \dot{x} &= \tilde{f}_0(x) + \sum_{i=1}^m v_i \tilde{f}_i(x) + \sum_{j=1}^r q_j g_j(x) \\ y &= h(x) \end{aligned}$$

où les champs de vecteurs  $\tilde{f}_i$  sont donnés par

$$\tilde{f}_0 = f_0 + \sum_{k=1}^m \alpha_k f_k \quad \text{et si } i \geq 1 \quad \tilde{f}_i = \sum_{k=1}^m \beta_{k,i} f_k. \quad (7.7)$$

Par analogie avec la définition que nous avons donnée des sous-espaces  $(A, B)$  invariants au chapitre 3, nous dirons que la distribution  $D$  est  $(f_0, f_1, \dots, f_m)$  invariante si il existe un feedback du type (7.6) telle que  $D$  soit invariante par les champs  $\tilde{f}_i$  donnés en (7.7) pour  $0 \leq i \leq m$ .

**THÉORÈME 7.5.** Le rejet des perturbations est possible localement pour le système (7.5) si et seulement s'il existe une distribution involutive  $D \subset \text{Ker}h$  qui est  $(f_0, f_1, \dots, f_m)$  invariante et contient les champs  $g_j$ .

Sous des conditions assez générales faisant intervenir des entiers dits *nombres caractéristiques* du système, on peut prouver l'existence d'une plus grande distribution  $D^*$   $(f_0, f_1, \dots, f_m)$  invariante contenue dans  $\text{Ker}h$ , que l'on peut calculer par un algorithme finitaire. La détermination d'un feedback (7.6) convenable devient alors explicite. Il convient de souligner que l'aspect local du résultat n'est pas véritablement dû à la définition de  $D^*$  mais plutôt à (7.6): en général, on ne pourra déterminer un tel feedback que sur le complémentaire d'une hypersurface qui est un ouvert très grand mais malheureusement non connexe. Des résultats analogues peuvent aussi être prouvés dans le cas lisse mais non analytique moyennant des hypothèses de constance de rang pour les distributions considérées.

Procédant d'une démarche voisine, l'une des idées les plus anciennes de la théorie est de linéariser un système non linéaire. Nous nous contenterons de discuter ceci dans un cas particulier mais représentatif. Il s'agit ici de linéarisation locale mais exacte, c'est à dire non pas d'approcher le système par un système

linéaire mais de le rendre véritablement linéaire par un changement de variable local adéquat –et non linéaire– dans l'espace d'état. On considère un système affine en la commande à une seule entrée

$$\dot{x} = f_0(x) + u f_1(x) \quad (7.8)$$

au voisinage d'un point d'équilibre  $x_0$  tel que  $f(x_0) = 0$ , et on se demande à quelle condition un changement de variable du type  $\xi = \phi(x)$ , où  $\phi$  est un difféomorphisme local défini sur un voisinage de  $x_0$  dans  $\mathbb{R}^n$  et tel que  $\phi(x_0) = 0$ , permet d'écrire

$$\dot{\xi} = A\xi + bu$$

où  $A$  et  $b$  sont des matrices constantes de taille  $n \times n$  et  $n \times 1$  respectivement, la paire  $(A, b)$  étant commandable. On dit dans ce cas que le système est localement linéarisable au voisinage de  $x_0$ .

**THÉORÈME 7.6.** Le système (7.8) est localement linéarisable au voisinage de  $x_0$  si et seulement si la distribution engendrée par les champs de vecteurs  $ad_{f_0}^k(f_1)$  lorsque  $k$  parcourt l'ensemble des entiers est de rang plein en  $x_0$  et vérifie en outre

$$[ad_{f_0}^k(f_1), ad_{f_0}^l(f_1)](x) = \sum_{i=0}^{\max(k,l)-1} \lambda_{i,k,l}(x) ad_{f_0}^i(f_1)(x)$$

pour certaines fonctions lisses  $\lambda_{i,k,l}$  définies au voisinage de  $x_0$ .

Ce résultat n'est pas exactement constructif en ce sens qu'il fait appel au théorème des fonctions implicites, mais la dérivée du changement de variable est, quant-à elle, connue. Des propriétés plus fortes peuvent être données dans le cas analytiques, toujours sous des hypothèses de finitude du type précédent.

### 7.3) Stabilisation.

La *stabilisation*, c'est à dire l'asservissement au voisinage d'une position d'équilibre par un feedback convenable, est un problème majeur en théorie du contrôle qui fait, à l'heure actuelle, l'objet d'une recherche intensive. D'une grande importance pratique, il est dans une certaine mesure plus abordable qu'un problème de commande général, du fait que l'état cible est toujours le même et qu'on ne cherche à l'atteindre qu'asymptotiquement. Néanmoins, il est le plus fréquemment fort délicat dans ses aspects globaux qui touchent à des questions fondamentales de la théorie des équations différentielles. Dans son énoncé *local*, la question s'énonce comme suit. Soit un système du type (7.1) et un *point d'équilibre*  $x_0$ , c'est à dire un point tel que  $f(x_0, 0) = 0$ . Trouver un feedback  $r(x)$  tel que le système bouclé par  $u = r(x)$  soit asymptotiquement stable dans un voisinage  $W$  de  $x_0$ , autrement dit tel que  $\lim_{t \rightarrow \infty} x(t) = x_0$  pourvu que la condition initiale  $x(0)$  appartienne à  $W$ . La version *globale* de la stabilisation s'énonce pareillement, à ceci près que le voisinage  $W$  est donné *a priori*. On souhaite idéalement obtenir  $W = \mathbb{R}^n$ , mais cela est simplement impossible dans certains cas où des raisons topologiques forcent la présence d'autres singularités que  $x_0$ . Examinons tout d'abord la question locale la plus simple: si l'approximation linéaire du premier ordre

$$\dot{\xi} = \frac{\partial f}{\partial x}(x_0, 0) \xi + \frac{\partial f}{\partial u}(x_0, 0) u$$

est exponentiellement stabilisable, et en particulier si ce système linéaire est commandable, il résulte du théorème de Poincaré mentionné dans l'introduction au chapitre 3 que la stabilisation locale est possible. Il est important de souligner aussi que si cette approximation linéaire est toujours exponentiellement instable, en particulier s'il existe des modes non commandables de partie réelle strictement positive, le système *n'est pas stabilisable*, une réponse négative souvent utile. Ce résultat, cependant, est de portée modeste si on observe que l'ouvert  $W$  dépend du feedback utilisé. C'est ici qu'un résultat de linéarisation comme le théorème 7.6 prend son sens car il permet, lorsqu'il s'applique, de s'affranchir de cet écueil et aussi de traiter certains cas dégénérés où l'approximation du premier ordre est stabilisable mais non exponentiellement. Hormis ces situations, tout problème de stabilisation devient un cas d'espèce susceptible d'être fort complexe. La seule approche générale repose sur l'exhibition d'une *fonction de Lyapunov*, c'est à dire d'une fonction lisse  $V : W \rightarrow \mathbb{R}$  qui est positive, ne s'annule qu'en  $x_0$ , qui est aussi *propre* c'est à dire que  $V(x)$  tend vers  $+\infty$

lorsque  $x$  quitte tout compact de  $W$ , et enfin qui décroît le long des trajectoires du système bouclé en un sens fort:

$$\sum_{i=1}^n \frac{\partial V}{\partial x_i} f^i(x, r(x)) < 0,$$

équation dans laquelle  $f^i$  désigne la  $i$ -ème composante de  $f$ . Des considérations physiques exprimant la dissipation de l'énergie sont à l'origine de la confection de la plupart des fonctions de Lyapunov en pratique. La méthode peut sembler bien particulière de prime abord, mais un théorème profond affirme en fait –du moins si  $r(x)$  est assez régulier, disons lisse sur  $\mathbb{R}^n - \{x_0\}$  pour simplifier– que l'existence d'un feedback stabilisant entraînera celle d'une fonction de Lyapunov, de façon peu explicite il est vrai. On peut objecter dans ces conditions que la difficulté est double: il faut tout à la fois construire  $V$  et  $r$  pour pouvoir conclure. Là encore, les apparences sont trompeuses comme le montre un résultat remarquable que nous n'énoncerons que dans le cas de systèmes affines en la commande. Soit donc un système du type (7.3), et désignons par  $f_k^i$  la  $i$ -ème composante du champ  $f_k$ . On posera

$$L_{f_k} V = \sum_{i=1}^n \frac{\partial V}{\partial x_i} f_k^i,$$

fonction qui s'appelle la *dérivée de Lie de  $V$  le long de  $f_k$* .

**THÉORÈME 7.7.** Si  $V : W \rightarrow \mathbb{R}$  est une fonction positive, propre et lisse, ne s'annulant qu'en  $x_0$ , et qui satisfait pour tout  $x \in W - \{x_0\}$

$$\inf_{u \in \mathbb{R}^m} \left( L_{f_0} V(x) + \sum_{k=1}^m u_k L_{f_k} V(x) \right) < 0. \quad (7.9)$$

Il existe alors un feedback  $r(x)$  stabilisant sur  $W$  pour le système (7.3) qui est lisse sur  $\mathbb{R}^n - \{x_0\}$ . De plus  $r(x)$  peut être choisi continu en  $x_0$  (resp. borné au voisinage de  $x_0$ ) pourvu que (7.9) puisse être obtenu avec des  $u_k$  arbitrairement petits (resp. bornés) lorsque  $x$  est assez voisin de  $x_0$ .

Ajoutons que, dans le cas où on a trouvé  $V$  satisfaisant les hypothèses du théorème, la détermination d'un feedback  $r$  convenable est explicite et même assez simple. Il faut noter que (7.9) exprime la possibilité de trouver des contrôles instantanés qui font décroître  $V$  le long de la trajectoire, montrant ainsi que le formalisme Hamiltonien du chapitre 6 sous-tend aussi le problème de la stabilisation. Il transparaît également de cet énoncé que  $r$  ne peut en général être choisi régulier en  $x_0$ , un problème que nous avons soigneusement passé sous silence et qui est lié à la dégénérescence en  $(x_0, 0)$  de l'approximation linéaire mentionnée plus haut. De nombreux travaux se donnent à l'heure actuelle pour tâche d'évaluer le degré de régularité qu'il est raisonnable d'exiger de  $r$ .

Avec un soupçon d'exagération, il ressort de ce qui précède que la stabilisation dans son acception globale peut être considérée comme l'art de déterminer des fonctions de Lyapunov. Il n'existe bien sûr aucune technique systématique pour ce faire. Mentionnons pour terminer que des recherches récentes concernant certains systèmes mécaniques modélisant des robots mobiles procèdent à l'élargissement de la classe des fonctions de Lyapunov, et donc facilitent la détermination de l'une d'entre elles, en utilisant, de façon peut-être inattendue, des feedback dépendant du temps et ce bien que le système soit stationnaire.

### 7.3) L'approche algébro-différentielle.

Il s'agit d'un abord de la théorie des systèmes inspiré du rôle que joue l'algèbre différentielle en théorie des équations algébro-différentielles, lequel vise à généraliser celui de l'arithmétique des corps en théorie des nombres ou plus vastement celui de l'algèbre commutative en géométrie algébrique. C'est dire qu'il s'agit d'un traitement fort abstrait qu'il serait vain de prétendre décrire substantiellement ici. Notre bref aperçu nécessitera tout d'abord quelques définitions. L'approche est, fondamentalement, *duale*: de la même manière qu'un sous-espace vectoriel peut être décrit au moyen des formes linéaires qu'il annule, qu'un nombre algébrique peut l'être par les relations qu'il satisfait ou une variété algébrique à travers les fonctions polynômes qui valent zéro en tous ses points, un système peut s'appréhender grâce aux relations différentielles qu'il engendre. Le lieu où placer de telles relations est un *corps différentiel*, c'est à dire un corps muni, outre des

opérations familières d'addition et de multiplication, d'une *dérivation*, c'est à dire d'une loi interne  $x \rightarrow \dot{x}$  vérifiant  $\overline{x_1 + x_2} = \dot{x}_1 + \dot{x}_2$  et  $\overline{x_1 x_2} = \dot{x}_1 x_2 + x_1 \dot{x}_2$ . On considère  $\mathbb{R}$  comme un corps différentiel en définissant la dérivation nulle: c'est le "corps des constantes". Si on a deux corps différentiels  $K_1$  et  $K_2$  et si  $K_1 \subset K_2$ , on dit que  $K_2$  est une *extension* de  $K_1$ . On se place une fois pour toutes dans une *extension universelle*  $K_u$  de  $\mathbb{R}$ , c'est à dire dans un corps différentiel contenant  $\mathbb{R}$  ainsi que toutes les extensions que nous aurons besoin de faire, et qui joue ici le rôle de la clôture algébrique en arithmétique ou du domaine universel en géométrie algébrique classique. Si  $K$  est un sous-corps et  $X$  un sous-ensemble de  $K_u$ , on note  $K \langle X \rangle$  la plus petite extension de  $K$  contenant  $X$ , et on dit que c'est l'extension de  $K$  engendrée par  $X$ . Lorsque on a une extension  $K_1 \subset K_2$ , on dit que  $x \in K_2$  est *différentiellement algébrique* sur  $K_1$  si  $x$  est solution d'une équation différentielle algébrique à coefficients dans  $K_1$ . Pour la différentiation usuelle, par exemple, la fonction  $\sin(t)$  est différentiellement algébrique sur  $\mathbb{R}$  car  $\sin^{(2)} + \sin = 0$ . La fonction  $g(t) = \text{Log}(t)$  qui vérifie  $t\dot{g} - 1 = 0$  n'est pas différentiellement algébrique sur  $\mathbb{R}$  mais elle l'est sur  $\mathbb{R}(t)$ , le corps différentiel des fractions rationnelles en une variable. Si tous les éléments de  $K_2$  sont différentiellement algébriques sur  $K_1$ , on dit simplement que l'extension est différentiellement algébrique. Lorsqu'un élément n'est pas différentiellement algébrique, on dit qu'il est *algébriquement transcendant* sur  $K_1$ . Des éléments  $x_\alpha$  de  $K_2$  sont dits différentiellement algébriquement indépendants sur  $K_1$  si aucun d'eux n'est différentiellement algébrique sur l'extension de  $K_1$  engendrée par les autres. Le nombre maximum d'éléments de  $K_2$  qui sont différentiellement algébriquement indépendants sur  $K_1$  est appelé le degré de transcendance différentiel de  $K_2$  sur  $K_1$ , que l'on note  $dtd(K_2/K_1)$ . Ce nombre peut être nul, lorsque l'extension  $K_2$  est différentiellement algébrique, ou au contraire infini. Lorsque  $K_2 = K_1 \langle x_1, x_2, \dots, x_d \rangle$  auquel cas on dit que l'extension  $K_2$  est de *type fini*,  $dtd(K_2/K_1)$  se trouve être égal au nombre maximum de  $x_i$  différentiellement algébriquement indépendants sur  $K_1$ . Ce nombre est donc nul si et seulement si chaque  $x_i$  est différentiellement algébrique sur  $K_1$ , et il est dans tous les cas inférieur ou égal à  $d$ . Si on considère une suite d'extensions  $K_1 \subset K_2 \subset K_3$ , l'additivité des degrés de transcendance est la propriété affirmant que  $dtd(K_3/K_2) + dtd(K_2/K_1) = dtd(K_3/K_1)$ . Appliquons à présent ces notions à un système dynamique dont  $u = (u_1, \dots, u_m)$  sont les entrées et  $y = (y_1, \dots, y_p)$  les sorties, ces quantités étant liées par des relations différentielles du type

$$y_i^{(n_i+1)} = f_i(u_1, \dots, u_m, \dots, u_1^{(n_i)}, \dots, u_m^{(n_i)}, y_1, \dots, y_p, \dots, y_1^{(n_i)}, \dots, y_p^{(n_i)})$$

où les  $f_i$  sont des fonctions rationnelles à coefficients fonctionnels. Définissons le corps de base  $k$  comme l'extension engendrée sur  $\mathbb{R}$  par les coefficients des  $f_i$ . Il découle à présent des équations du système que  $k \langle u, y \rangle$  est une extension différentiellement algébrique de  $k \langle u \rangle$ . Ceci peut aussi bien être retourné en une définition du système lui-même: *un système est une suite d'extensions  $k \subset k \langle u \rangle \subset k \langle u, y \rangle$  où  $k$  est un corps différentiel, usuellement une extension de  $\mathbb{R}$  par un certains nombre de fonctions,  $k \langle u \rangle$  une extension de type fini, et  $k \langle u, y \rangle$  une extension de type fini différentiellement algébrique de  $k \langle u \rangle$ . Afin d'atténuer l'impression d'inutile abstraction qui pourrait s'emparer du lecteur à ce stade, illustrons ces définitions avec le problème à la fois pratique et fondamental de l'*inversion à gauche* du système, c'est à dire la possibilité de calculer l'entrée en intégrant des équations différentielles dépendant de la sortie et des fonctions de base du corps  $k$ . Ceci revient à demander si  $k \langle u, y \rangle$  est différentiellement algébrique sur  $k \langle y \rangle$ . Le résultat suivant est alors une conséquence immédiate de l'additivité des degrés de transcendance: THÉORÈME 7.8. Le système est inversible à gauche si et seulement si  $dtd(k \langle u \rangle / k) = dtd(k \langle y \rangle / k)$ .*

L'entier  $dtdk \langle y \rangle / k$  est appelé rang différentiel de sortie du système, et l'énoncé ci-dessus affirme que le système est inversible à gauche si ce rang est égal au nombre d'entrées indépendantes. On prouverait de même que le système est inversible à droite si le rang différentiel de sortie est égal au nombre de sorties. Comme il n'est pas de moyen systématique à l'heure actuelle (sauf dans le cas d'une seule entrée et d'une seule sortie avec des équations de degré 1) de tester l'algébricité différentielle sur un sous-corps, ceci ne constitue pas un test constructif de l'inversibilité en général mais il faut saluer la netteté du résultat et l'adéquation des concepts. En fait, l'algèbre différentielle est à même de donner des énoncés synthétiques et conceptuellement clarificateurs concernant la plupart des questions structurelles en théorie des systèmes, y compris réalisation, découplage et linéarisation. Le cadre des équations implicites, très courantes en pratique, devient le cadre naturel, et la généralisation à certains types d'équations non algébriques est également possible, Il faut cependant garder présent à l'esprit un certain nombre de points. Tout d'abord, on ne connaît la solution d'une équation différentielle que si on en sait la condition initiale. Ainsi, la connaissance des relations de dépendance différentielle de  $u$  sur  $y$  dont le théorème 7.8 affirme, sous certaines conditions, l'existence ne

suffit pas *stricto sensu* pour inverser le système car il faut disposer d'une sorte de stabilité pour atténuer l'effet de cette condition initiale avec le temps. Cette notion de stabilité amène naturellement à considérer une difficulté générale de cette approche à savoir qu'il ne suffit pas de donner un nom à la solution d'une équation différentielle pour savoir comment elle se comporte et tout d'abord dans quel espace fonctionnel elle peut être représentée. En particulier, les équations différentielles que l'on manipule ici sont la plupart du temps singulières car leur coefficients, qui appartiennent à un corps formel de fonctions, voient en général leur dénominateur s'annuler en certains points. On se trouve là dans une situation vaguement analogue, quoique bien plus compliquée encore, à celle qui prévaut en géométrie algébrique où, lorsqu'on veut étudier une variété plongée, disons dans  $\mathcal{C}^N$  ou mieux encore dans  $\mathbb{R}^n$ , la topologie au voisinage des points non-génériques peut se révéler très complexe. Les notions de réalité et de valuation qui semblent nécessaires à l'abord de ces questions topologiques ne paraissent d'ailleurs pas avoir été clairement dégagées pour l'instant en ce qui concerne les systèmes. Bien que l'approche algébro-différentielle soit trop récente pour qu'on puisse prétendre analyser définitivement une contribution d'ores et déjà importante, il est cependant probable que de nombreux outils manquent encore pour établir un pont entre l'algèbre et la géométrie des systèmes dynamiques.

## 8. Filtrage Non-Linéaire

Considérons le cas général où l'état du système est décrit par une équation différentielle non-linéaire bruitée

$$dX_t = f(X_t) dt + g(X_t) dV_t$$

et où l'on dispose seulement d'observations non-linéaires et bruitées sur l'état du système

$$dY_t = h(X_t) dt + dW_t .$$

On dénote par  $M$  et  $N$  respectivement, la matrice de covariance des bruits (processus de Wiener ou "marche aléatoire")  $\{V_t, t \geq 0\}$  et  $\{W_t, t \geq 0\}$  supposés gaussiens indépendants. Comme dans le cas du filtre de Kalman, le sens précis à donner aux équations ci-dessus, le calcul de Ito, est un peu délicat mathématiquement, et ceci d'autant plus que  $g$  peut dépendre de  $X$ . Nous ne nous étendrons pas sur cet aspect technique.

Une approche naturelle, et utilisée avec succès dans de nombreuses applications, consiste à linéariser les équations du système autour de la valeur courante de l'estimateur, et à définir cet estimateur comme le filtre de Kalman pour le système linéarisé. On obtient ainsi les équations du filtre de Kalman étendu

$$d\hat{X}_t = f(\hat{X}_t) dt + K_t [dY_t - h(\hat{X}_t) dt]$$

$$\dot{\Sigma}_t = F_t \Sigma_t + \Sigma_t F_t' + G_t M G_t' - \Sigma_t H_t N^{-1} H_t' \Sigma_t$$

avec le gain de Kalman  $K_t = \Sigma_t H_t' N^{-1}$  et les coefficients du système linéarisé

$$F_t = \frac{\partial f}{\partial x}(\hat{X}_t) \quad H_t = \frac{\partial h}{\partial x}(\hat{X}_t) \quad G_t = g(\hat{X}_t) .$$

L'efficacité du filtre de Kalman étendu, constatée dans de nombreuses applications, peut être démontrée mathématiquement dans un cadre asymptotique, sous des hypothèses de détectabilité. Il existe néanmoins certaines situations, caractéristiques des systèmes non-linéaires, comme l'existence de plusieurs bassins d'attraction pour l'équation d'état, ou la non-observabilité du système, qui combinées avec un manque d'information sur la condition initiale, rendent le filtre de Kalman étendu peu efficace. Il est parfois possible de remédier à cette situation, en couplant par exemple une batterie de filtres de Kalman étendus opérant en parallèle, et des règles de décision statistique permettant de choisir entre ces différents estimateurs.

On peut aussi, quand la dimension de l'espace d'état le permet, adopter une approche de plus haut niveau, où l'estimateur est représenté par la densité de probabilité conditionnelle de l'état du système, sachant les observations :  $p_t(x)$  L'évolution de cette densité est décrite de façon récursive par une équation aux dérivées partielles stochastique, appelée équation de Zakai:

$$dp_t = L^* p_t dt + p_t h' N^{-1} dY_t,$$

où  $L^*$  est l'adjoint de l'opérateur elliptique  $L$  qui à toute fonction  $q(x)$  fait correspondre la fonction

$$Lq = \frac{1}{2} \operatorname{tr} \left( g' \frac{\partial^2 q}{\partial x^2} g M \right) + \frac{\partial q}{\partial x} f$$

(Écrire le dual explicitement est sans intérêt car dans les procédures numériques, cela revient simplement à transposer la matrice de l'opérateur). On constate souvent que les situations où le filtre de Kalman étendu est inefficace correspondent, au niveau de la densité conditionnelle, à l'apparition de plusieurs pics. Il s'agit en général d'une situation transitoire, qui ne se résout qu'après que le traitement des mesures suivantes ait permis d'éliminer les pics parasites. On comprend ainsi qu'un filtre de Kalman étendu unique, pour lequel le choix entre les différentes options doit être instantané, puisse ne pas être efficace dans une telle situation.

### Bibliographie

Comme nous l'avons indiqué, ceci est une bibliographie très partielle, indiquant quelques ouvrages permettant de compléter utilement cet article, et choisis surtout en fonction de leur caractère didactique.

- K.J. Aström and B. Wittenmark: Adaptive Control, Addison Wesley, 1989
- T. Başar and P. Bernhard:  $H^\infty$  Optimal Control and Related Minimax Design Problems: a Dynamic Game Approach, Birkhauser, 1991.
- A. Bensoussan: Stochastic Control by Functional Analysis Methods, North Holland, 1982
- P. Faure et M. Robin: Eléments d'Automatique, deuxième édition, Dunod, 1984
- W. Fleming and R. Rishel: Deterministic and Stochastic Optimal Control, Springer, 1975
- P.A. Fuhrmann: Linear Systems and Operators in Hilbert Space, Mac Graw Hill, 1981.
- A. Isidori: Nonlinear Control Systems: An Introduction, 2me édition, Springer, 1989.
- B.D.O. Anderson, R.R. Bitmead, C.R. Johnson, P.V. Kokotovic, R.L. Kosut, I.M.Y. Mareels, L. Praly, and B.D. Riedle: Stability of Adaptive Systems, the MIT Press, 1986.
- B. Francis: A Course in  $\mathcal{H}_\infty$  Control Theory, Lecture notes in control and information sciences 88, Springer, 1987.
- V. Kučera: Discrete Linear Control, Wiley, 1979
- J.L. Lions: Contrôle des systèmes gouvernés par des équations aux dérivées partielles, Dunod, 1968
- W.M. Wonham, Linear Multivariable Control: A Geometric Approach, 2nd edition, Springer, 1979