

UNIVERSITÉ D'ALGER  
ECOLE NATIONALE POLYTECHNIQUE  
DEPARTEMENT ÉLECTRICITÉ

9/74

المدرسة الوطنية للعلوم الهندسية  
المكتبة  
ECOLE NATIONALE POLYTECHNIQUE  
BIBLIOTHÈQUE

3ea

# PROJET DE FIN D'ETUDES

المدرسة الوطنية للعلوم الهندسية  
المكتبة  
ECOLE NATIONALE POLYTECHNIQUE  
BIBLIOTHÈQUE

## IDENTIFICATION DYNAMIQUE D'UNE MACHINE A PAPIER

PROPOSÉ PAR :  
G. BETHOUX

ETUDIÉ PAR :  
S. GUERMAH et N. MAAFI

Nous tenons à remercier très sincèrement Monsieur Guy BETHOUX , de la Direction des Recherches d'ALSTHOM de Grenoble, qui nous a guidés dans l'élaboration de ce projet.

L'aide qu'il nous a apportée notamment dans la conception, assez délicate, du chapitre VII et les conseils qu'il nous a prodigués tout au long de ce travail nous ont été très profitables . Qu'il trouve ici l'expression de notre profonde reconnaissance.

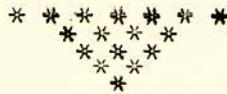
Nos remerciements vont également aux membres du jury qui nous ont fait honneur de leur présence.

Nous n'oublierons pas l'équipe du Centre de Calcul pour sa contribution à la réalisation de ce projet.

Pour l'assistance qu'il nous a prêtée dans la confection de ce fascicule, nous adressons nos vifs remerciements au service FORMATION de SONELEC .

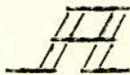
Enfin, que tous les professeurs de l'Ecole Nationale Polytechnique trouvent ici l'expression de notre gratitude pour l'enseignement qu'ils nous ont dispensé durant notre formation .

T A B L E   D E S   M A T I E R E S



	page
A V A N T * P R O P O S	I
I N T R O D U C T I O N	VI
CHAPITRE I : POSITION DU PROBLEME	1
I-1 Choix du modèle et du critère	1
I-2 Les méthodes	4
CHAPITRE II : ETUDE DU MODELE	8
II-1 Analyse du modèle	8
II-2 Les résidus	11
II-3 Le retard pur	12
II-4 Représentation de la fonction de transfert échantillonnée - Gain statique	13
II-5 Problèmes d'échantillonnage	14
II-6 Stabilité	15
CHAPITRE III : METHODE DES MOINDRES CARRÉS SIMPLES	17
A) OPTIQUE TEMPS DIFFERE	
III-1 Le critère à minimiser et la solution	18
B) OPTIQUE TEMPS REEL	
CHAPITRE IV : METHODE DE LA MATRICE INSTRUMENTALE	28
A) OPTIQUE TEMPS DIFFERE	
IV-1 Principe de la méthode	28
IV-2 Quelques exemples de matrices instrumentales	29
IV-3 Estimation de la variance des résidus	33

	page
B) OPTIQUE TEMPS REEL	
IV-4 Algorithme d'estimation des paramètres	34
IV-5 Estimation des résidus en temps réel	35
CHAPITRE V : METHODE DES MOINDRES CARRÉS GENERALISÉS	
A) OPTIQUE TEMPS DIFFÉRE	
V-1 Méthodes de filtrage	37
V-2 Modèle paramétrique élargi	44
B) OPTIQUE TEMPS REEL	
V-3 Filtre de KALMAN appliqué à l'identification	47
V-4 Autres méthodes des MCG en temps réel	50
CHAPITRE VI : LES SYSTEMES ADAPTATIFS ECHANTILLONNES ET LEUR UTILISATION POUR L'IDENTIFICATION DES PROCEDES PHYSIQUES	
VI-1 Rappel des concepts de base	54
VI-2 Application à l'identification	56
ANNEXE : DETERMINATION DU CORRECTEUR LINEAIRE	61
CHAPITRE VII : RESULTATS EXPERIMENTAUX	
VII-1 Description de la machine à papier	64
VII-2 Techniques de mise en oeuvre des méthodes d'identification	65
VII-3 Résultats d'identification des paramètres du transfert débit pâte-humidité de la machine à papier	72
VII-4 Les modèles obtenus et leurs performances	73
C O N C L U S I O N	
	76
B I B L I O G R A P H I E	
	78



V A N T - P R O P O S

Avant la seconde guerre mondiale, le problème de la commande des processus et surtout celui de la commande optimale, était considéré comme un art. L'automatique avait en effet commencé d'une façon très empirique; les mathématiques n'y avaient pas une grande place. Cette absence de théorie était d'ûe surtout au fait que la technologie ne l'exigeait pas. Il s'agissait dans la plupart des cas de boucler un système par l'intermédiaire d'un correcteur de façon à obtenir une sortie adéquate avec un temps de réponse aussi petit que possible, tout en assurant la stabilité du système.

L'apparition de systèmes de plus en plus complexes fit sentir le besoin d'utiliser plus largement les mathématiques et de créer ou d'adapter des théories et des moyens d'étude nouveaux. L'outil fondamental en a été la transformée de LAPLACE, et la notion de fonction de transfert a apporté une grande unité. Les théories de BODE, de NYQUIST (analyse harmonique) et celle d'EVANS (méthode des zéros et des pôles) sont à présent classiques.

Avec l'évolution des technologies, l'apparition des systèmes à informations discrètes et l'utilisation de calculateurs dans les chaînes de commande, une théorie s'est développée pour les systèmes échantillonnés, linéaires, monovariables, avec la transformée en Z notamment.

Le rôle et la place que tiennent aujourd'hui les calculateurs numériques dans la commande automatique deviennent donc prépondérants. On peut se poser la question de savoir pourquoi utiliser ces derniers, de préférence aux systèmes analogiques qui présentent pourtant de réels avantages : ils travaillent essentiellement en "visuel" et leur réalisation se réduit à l'interconnexion de composants, de façon directe. Des unités fonctionnelles (engrenages, potentiomètres, ...) assument dans ces systèmes une fonction propre de façon à obtenir une fonction de transfert qui ait les paramètres désirés. De plus, la caractéristique temps réel des techniques analogiques est très intéressante et leur nature continue les rend plus adaptées pour les asservissements en boucle fermée. Ainsi, quand la fonction de transfert est compliquée, un dispositif analogique peut répondre plus vite qu'un système digital programmé, à échantillonnage périodique.

Néanmoins, en termes de précision et de fiabilité, ces systèmes restent limités par la précision des composants qui les constituent. Des raisons économiques imposent souvent des restrictions sur cette précision; il y a ici une considération pratique, à savoir combien de temps et d'argent peut-on consacrer au développement de composants ultra-précis pour un système analogique donné. En pratique, les précisions dans le cas de systèmes complexes sont limitées à 1/1000e, et certaines opérations sont sujettes à des imperfections (telle la tension d'offset dans les multiplieurs); cette approche a donc les avantages d'une grande rapidité et une bonne aptitude en temps réel, son point faible étant sa limitation en précision et dans l'exécution de certaines opérations que l'on n'a pas encore pu mettre sous forme analogique, comme par exemple certaines fonctions mathématiques non linéaires.

Si l'on examine le cas des systèmes digitaux, qui traitent les données sous forme discrète, ils ont à leur actif la largeur des tolérances pour les composants auxquels on ne demande pas une précision très poussée. Un degré quelconque de précision peut être atteint et maintenu pour le fonctionnement de ces systèmes grâce à un choix judicieux du codage des nombres: pour le code binaire naturel, on obtient 1/1000e avec 10 caractères, 1/2000e avec 11 caractères, ... Par ailleurs, le calculateur est d'un emploi suffisamment souple pour exécuter tout type d'opération mathématique et possède une grande fiabilité. Il a l'inconvénient d'être plus complexe et d'un prix relativement plus élevé à l'acquisition (si l'on inclut tous les organes annexes). L'obstacle principal réside toutefois dans la suite des opérations d'échantillonnage, de transfert des données en mémoire, de leur traitement suivant les instructions d'un programme, de transfert des résultats en mémoire, de génération de données de contrôle, etc... Ces opérations peuvent demander beaucoup de temps. Par conséquent, la vitesse d'échantillonnage doit être limitée et il en résulte une perte d'informations. Pour cette même raison, la commande élaborée n'est pas en synchronisme parfait avec la réception des données.

Une tendance s'est dessinée pour l'emploi des deux approches à la fois, afin de réunir leurs avantages :

-prix de revient plus bas (grâce -non pas à un calculateur complet- mais à des éléments digitaux incorporés à l'ensemble),

-contrôle plus serré, plus grande rapidité qu'avec les systèmes programmés,

- complexité moindre que celle des unités programmées, seules,
- aptitude à traiter des données analogiques (qui sont ainsi dans leur forme normale en général) , ces données peuvent être converties en numérique dans ce même ensemble.

Un tel dispositif, dit hybride, n'offre pourtant pas toujours la meilleure solution pour un problème de calcul et son application n'est utile et économiquement possible que si de nouveaux composants, très performants, sont construits.

Depuis quelques années, les progrès accomplis dans le domaine des calculateurs, notamment du point de vue rapidité, et leur prolifération sur le marché (rendant leur prix de plus en plus compétitifs), facilitent beaucoup le choix qui s'est imposé au départ. L'emploi de micro-calculateurs se répand de plus en plus dans l'industrie, où leurs performances permettent largement d'atteindre les résultats que l'on peut escompter.

Les unités industrielles (laminoirs, centrales thermiques, unités de production chimiques et pétrolières ...) sont en général équipées d'un grand nombre de boucles d'asservissement dont les valeurs de consigne étaient et sont encore dans beaucoup de cas fixées par l'ingénieur responsable de la marche de l'unité. Cet ingénieur cherche à en optimiser le fonctionnement suivant des critères qualitatifs, pas toujours très bien définis. Cette optimisation dépend des capacités et de l'expérience de l'ingénieur, de la complexité de l'installation.

L'évolution actuelle est d'essayer de faire de l'optimisation quantitative à partir de critères bien définis. Ceci n'est possible qu'en utilisant un ordinateur, numérique dans la plupart des cas, qui aidera dans un premier stade et remplacera dans un second, l'homme dans sa fonction de conducteur de l'unité. Les tâches assignées au ordinateur dans ce cadre sont de deux sortes :

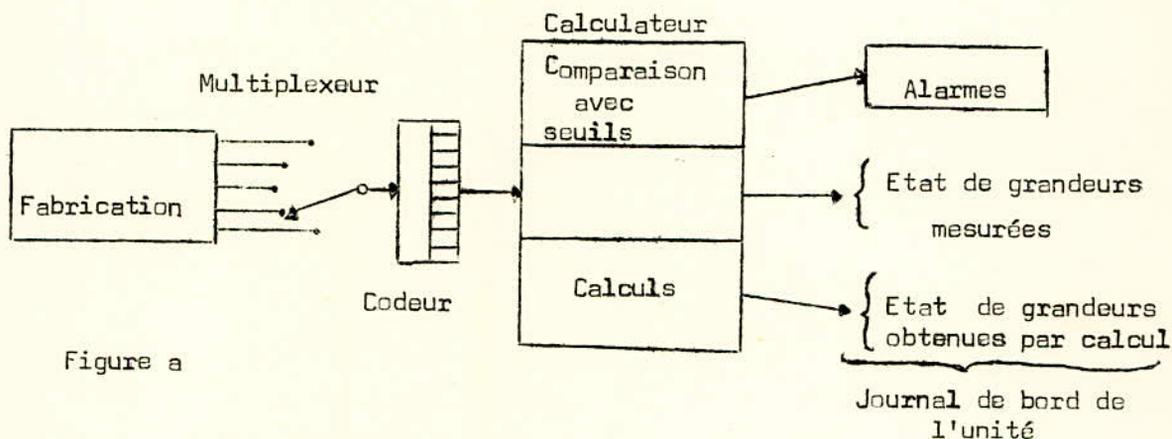
a) Fonction de surveillance et de collecte d'informations

L'ensemble des opérations de mesures (tests, relevés de paramètres, ou de variables importantes du processus à automatiser), nécessitant en général l'installation de capteurs et appareils de mesure, peut être assuré par le ordinateur qui est alors utilisé comme un dispositif de collecte d'information. Les valeurs imprimées par le ordinateur peuvent être :

- soit mesurées directement au moyen de capteurs,
- soit élaborées par le ordinateur en fonction de grandeurs mesurées.

Les informations reçues par le calculateur peuvent être aussi utilisées pour détecter par comparaison avec des seuils préaffichés, des conditions de fonctionnement anormal et déclencher alors des alarmes. Le calculateur assure dans ce cas une fonction de surveillance et peut être utilisé pour enregistrer les séquences des événements ayant conduit au fonctionnement anormal constaté ce qui permettra d'en connaître avec précision les causes.

Dans ce type de fonctionnement, il faut remarquer que le calculateur reçoit seulement des informations du processus auquel il est couplé sans réagir sur ce dernier: il n'y a pas de bouclage du calculateur sur l'unité. Les avantages de ce type d'utilisation sont intéressants puisqu'ils peuvent aboutir à des gains qualitatifs et quantitatifs sur la production. (cf. figure a)



#### b) Fonction d'optimisation

A partir des données obtenues et des mesures complémentaires, on peut, pour certaines classes de procédés, élaborer à l'aide du calculateur un modèle mathématique du processus étudié. Dans ce cas, le calculateur peut assurer une fonction d'optimisation, c'est-à-dire obtenir les consignes les plus favorables au fonctionnement optimal de l'unité par des calculs souvent trop longs ou trop complexes pour l'homme.

Le calculateur prend alors une part active à la commande de l'unité de fabrication qu'il se contentait jusque là de contrôler. (cf. figure b)

En résumé, l'introduction d'un calculateur numérique dans la commande d'un processus pose trois problèmes importants :

i) Le problème de l'identification de ce processus, qui permettra d'obtenir un modèle mathématique, problème dont la solution est facilitée par la présence du calculateur.

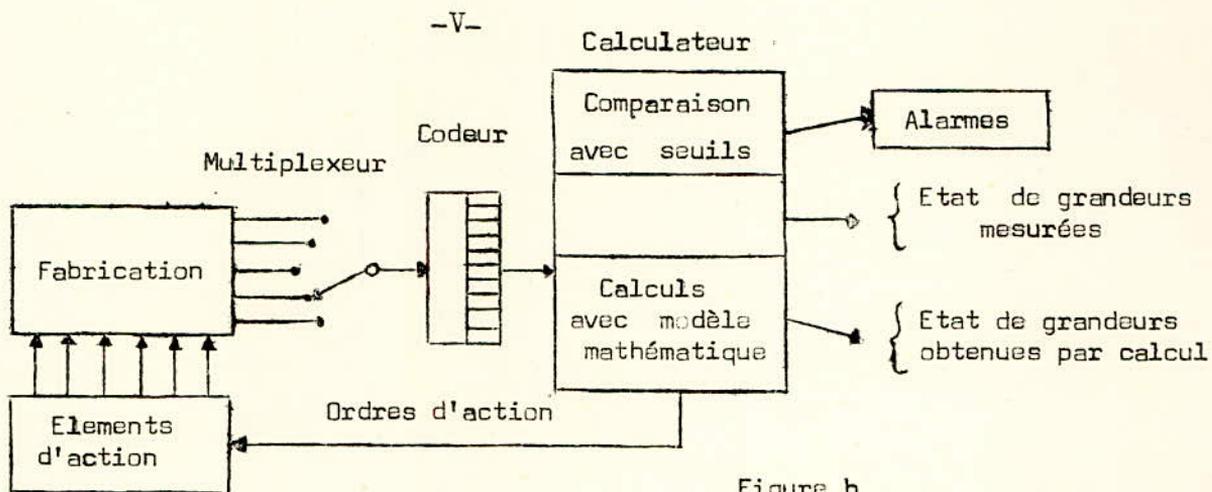
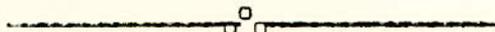


Figure b

ii) Le problème du choix des méthodes d'optimisation qui permettront d'obtenir le fonctionnement optimal souhaité.

iii) Le problème de la commande effective du processus par le calculateur (c'est-à-dire l'ensemble des problèmes liés à l'introduction d'un calculateur dans une chaîne de commande) .

Nous nous sommes personnellement intéressés dans l'étude qui va suivre , au problème de l'identification .



## II INTRODUCTION

L'intérêt croissant porté aux systèmes asservis de plus en plus complexes et variés a incité les automaticiens à utiliser les progrès de la technologie dans le domaine de la collecte et du traitement des informations (calculateurs, capteurs,...), et à appliquer ces nouvelles possibilités à l'automatisation. En plus de la stabilité et de la rapidité du système à suivre une consigne, on cherche à calculer des commandes qui en optimisent le fonctionnement au sens d'un critère dont le choix est laissé à l'utilisateur. Ce critère essaie de tenir compte de l'identité entre sortie et consigne, mais aussi de l'énergie dépensée et de la sécurité vis à vis des contraintes.

Le calcul de ces commandes nécessite la connaissance des équations qui rendent compte du fonctionnement du procédé. Les équations statiques ou dynamiques selon les buts de cette commande. La mise au point des techniques qui permettent de trouver ces équations est ce que l'on appelle identifier le procédé. Autrement dit l'identification est le passage d'un ensemble concret à un ensemble abstrait et se définit comme la recherche d'un modèle mathématique qui représente, avec une précision suffisante, le comportement du système compte tenu de son état et des actions qui le sollicitent. Cela constitue la première étape de la résolution du problème de la commande des processus; son importance est capitale, puisque toute la suite de l'étude théorique dépendra de la validité du modèle mathématique retenu.

Pour effectuer cette identification, les moyens dont nous disposons sont :

- Les lois fondamentales de la physique qui permettent, en général, d'avoir une idée de la structure du modèle mathématique ,

- L'expérimentation, parfois difficile, permettant de préciser notamment les valeurs numériques,

La précision est a priori la qualité essentielle d'une identification, c'est-à-dire que le modèle, soumis aux mêmes entrées que le processus réel, doit représenter fidèlement l'évolution des sorties réelles. Mais pour être utilisable, ce modèle ne doit pas être trop complexe; plus il est simple, plus l'étude théorique sera facile.

Il faut toujours se rappeler que le modèle mathématique n'est qu'un moyen, qui servira au calcul des commandes. Sa simplicité permet d'alléger l'étude théorique et d'économiser le temps de calcul quand il s'agit d'inclure un calculateur dans la chaîne de commande. Par ailleurs, le domaine de fonctionnement qui nous intéresse est souvent restreint par les contraintes ou éventuellement par une idée que l'on peut avoir de la forme des commandes à appliquer. Ainsi, une approximation linéaire peut être très valable pour représenter un fonctionnement vérifiant les contraintes, bien que des non-linéarités importantes apparaissent si on sort de ce domaine.

En conclusion; on peut dire que l'établissement du modèle mathématique-qui est toujours une approximation de la réalité-résulte d'un compromis entre la précision et la simplicité et qu'il faut tenir <sup>compte</sup> dès ce stade du but à atteindre.

## CHAPITRE I

### POSITION DU PROBLEME

#### I-1 CHOIX DU MODELE ET DU CRITERE

ZADEH a donné la définition suivante de l'identification : "l'identification est la détermination, basée sur la connaissance des entrées et des sorties d'un procédé, d'un modèle, appartenant à une classe spécifiée, équivalent au procédé".

Deux notions essentielles apparaissent dans cette définition: celle de classe de modèle et celle d'équivalence.

#### I-1.1 Caractérisation

Il est fondamental de se fixer la classe de modèle. Ce choix est appelé caractérisation. Il s'agit de donner au modèle une forme bien déterminée.

a) L'hypothèse de départ d'une caractérisation est la linéarité du système: elle est généralement vérifiée pour les systèmes industriels qui travaillent par petites variations autour du point de fonctionnement nominal. Quand on parle de linéarité en identification, on distingue deux types:

-linéarité par rapport aux variables d'entrée-sortie.

-linéarité par rapport aux paramètres: les équations du système sont alors linéaires par rapport aux paramètres mais peuvent être non linéaires par rapport aux entrées-sorties.

b) Il convient ensuite de se donner le type de modèle et sa représentation (modèle continu ou discret, représentation paramétrique ou non paramétrique.)

Pour un modèle paramétrique continu, on cherchera à identifier les coefficients d'une équation différentielle, comme

par exemple:

$$y''(t) + a_1 y'(t) + a_2 y(t) = b_1 u'(t) + b_2 u(t)$$

ou bien d'une équation aux différences, dans le cas discret, telle que:

$$y(k-2) + a_1 y(k-1) + a_2 y(k) = b_1 u(k-1) + b_2 u(k)$$

ou encore des fonctions de transfert, telles que:

$$\frac{K}{1 + T_p} \quad ; \quad \frac{G}{Z - a}$$

Dans les modèles non paramétriques, on classe les réponses impulsionnelles  $h(t)$ , telles que:

$$y(t) = \int_0^{\infty} h(\tau) u(t-\tau) d\tau \quad \begin{cases} u(t-\tau) = 0 \\ \text{pour } t < \tau \end{cases}$$

ainsi que les séquences de pondération:

$$y(k) = \sum_{j=1}^{\infty} H(j) u(k-j+1) \quad \begin{cases} u(k-j+1) = 0 \\ \text{pour } k-j+1 < 0 \end{cases}$$

Chacune de ces représentations a bien sûr des avantages et des inconvénients qui lui sont propres.

La caractérisation est une phase délicate de l'identification, elle dépend :

- des connaissances a priori que l'on peut avoir sur le système,
- des moyens de calcul dont on dispose,
- du but du modèle lui même.

### I-12 Equivalence

Cette notion s'exprime par l'intermédiaire d'un critère qu'il s'agit de minimiser; c'est l'identification proprement dite: il s'agit d'affecter des valeurs numériques aux paramètres choisis.

Il existe un grand nombre de critères possibles. Ces critères sont l'expression mathématique de l'écart entre le système et le modèle. Si on se place dans l'espace paramétrique,

espace des paramètres  $\underline{\theta}_i$  à déterminer, on peut définir dans cet espace une distance  $D$  du point  $M$  représentant le modèle au point représentant le procédé,  $P$ . Cette distance est définie par le critère. On rencontre souvent ces trois types de distances:

-Distance de structure: c'est la norme, quelconque, de la différence entre le vecteur (ou la matrice) des paramètres du modèle et le vecteur (ou la matrice) des paramètres du procédé.

$$\Delta \underline{\theta} = \underline{\theta}_m - \underline{\theta}_p \quad D = f(\Delta \underline{\theta})$$

exemple :  $D = \Delta \underline{\theta}^T \cdot A \cdot \Delta \underline{\theta}$   $A$  : matrice définie positive  
(critère quadratique)

-Distance d'état : c'est une norme quelconque de la différence entre l'état (ou les états) du procédé et celui (ou ceux) du modèle.

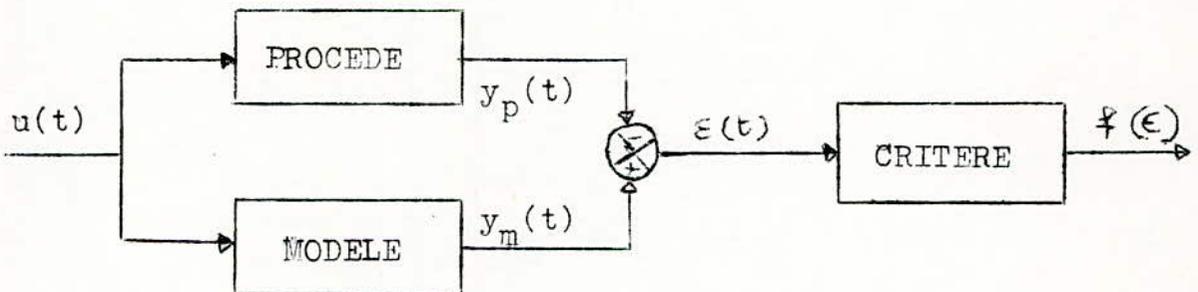
Si  $y_m(t)$  est la sortie du modèle et  $y_p(t)$  la sortie du procédé soumis à la même entrée  $u(t)$ , on pose:

$$\epsilon(t) = y_m(t) - y_p(t)$$

$$D = f(\epsilon)$$

$y_m(t)$  est calculé ici à partir de  $u(t)$  et  $\underline{\theta}$  :

$$y_m(t) = g[u(t), \underline{\theta}]$$



On peut prendre, par exemple:

-un critère absolu:  $D = \sum_{t=1}^N |\epsilon(t)|$

-un critère quadratique:  $D = \sum_{t=1}^N \epsilon^2(t)$

Cette distance d'état est particulièrement significative

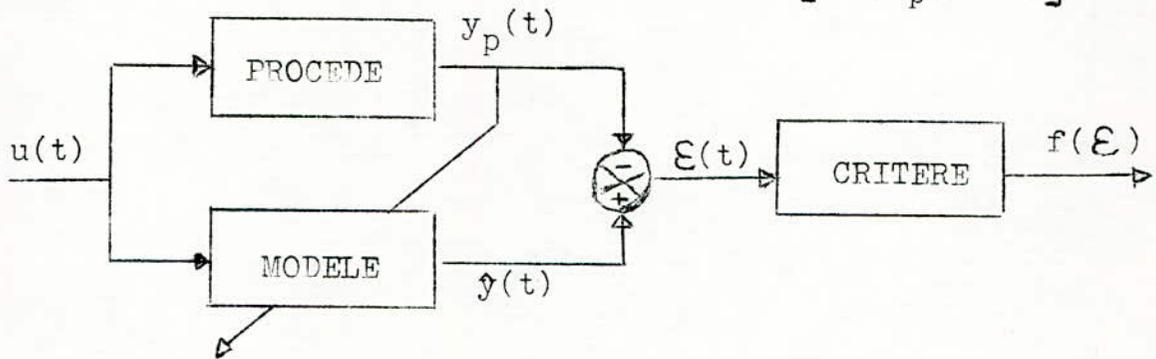
puisque'elle consiste à choisir le vecteur  $\underline{\theta}$  qui fait le mieux coïncider la sortie du système et celle du modèle.

Erreur de prédiction : on la définit par:

$$\varepsilon(t) = \hat{y}(t) - y_p(t) \quad ; \quad D = f(\varepsilon)$$

où  $\hat{y}(t)$  représente la sortie prédite par le modèle à partir de toutes les mesures précédentes. Ce calcul fait intervenir les mesures de l'entrée et de la sortie du système.

$$\hat{y}(t) = g[u(t), y_p(t-1), \underline{\theta}]$$



Une fois le critère choisi, il faut le minimiser. Pour cela, on dispose de nombreuses méthodes.

## I-2 LES METHODES

La minimisation du critère est un problème mathématique. Nous allons décrire les moyens utilisés en identification pour la résolution de ce problème.

### I-21 - Le programmation non linéaire (P.N.L.)

Nous n'entrerons pas dans les détails de cette théorie, l'objet principal de notre étude étant les méthodes statistiques. Le PNL est un problème d'optimisation paramétrique dans lequel il s'agit de minimiser un critère par l'une des méthodes suivantes:

- méthode du gradient, la plus courante,
- méthode de NEWTON-RAPHSON ,
- méthode de GAUSS-NEWTON ,
- méthode de l'approximation stochastique .

Ces méthodes sont basées sur le calcul du gradient du critère par rapport à  $\underline{\theta}$  :

$$C_{\theta}(\underline{\theta}_i) = \left[ \frac{\partial C}{\partial \theta_1} \quad \frac{\partial C}{\partial \theta_2} \quad \dots \quad \frac{\partial C}{\partial \theta_n} \right] (\underline{\theta}_i)$$

et ne diffèrent entre elles que par l'ordre du développement de  $C(\underline{\theta}_i)$ . Elles sont maintenant classiques et on montre que leur emploi nécessite une certaine prudence, notamment dans le choix des entrées.

### I-22 Les méthodes statistiques

Puisque les mesures prélevées sur le système sont entachées de bruit, la minimisation du critère consiste à considérer l'identification comme un traitement de l'information, donc comme un problème statistique. Les mesures sont traitées comme des variables aléatoires dont les propriétés dépendent essentiellement de la nature du bruit. Les nombreuses méthodes permettant de trouver le vecteur  $\underline{\theta}$  des paramètres du modèle, diffèrent entre elles par le nombre d'hypothèses formulées, en particulier sur la nature du bruit. Elles permettent d'attribuer une variance aux paramètres, donc d'avoir une indication sur la précision de l'identification.

La détermination des paramètres par des méthodes statistiques est basée sur l'hypothèse suivante : les mesures effectuées sur le procédé, résumées dans le vecteur  $\underline{Y}$  sont fonction des paramètres  $\underline{\theta}_p$  du système et de variables déterministes  $X$ , et sont entachées d'un bruit  $\underline{b}$ . Le modèle prend alors la forme suivante:

$$\underline{Y} = f(\underline{\theta}_p, X) + \underline{b}$$

A partir des mesures obtenues, on cherche une estimation  $\hat{\underline{\theta}}$  du vecteur  $\underline{\theta}_p$ ;

Parmi les méthodes statistiques, la méthode de BAYES est la plus élaborée. Elle minimise l'espérance mathématique d'un critère, celui-ci étant fonction de  $\underline{\theta}$  et  $\underline{\theta}_p$  où  $\underline{\theta}_p$  est

considéré comme un vecteur aléatoire. Nous ne l'étudierons pas car elle nécessite une grande connaissance des lois de probabilité des variables intervenant dans le modèle et présente peu d'intérêt pour l'identification.

La méthode du maximum de vraisemblance, considérant, elle, le vecteur  $\underline{\theta}$  comme déterministe, est plus accessible. Elle rend la probabilité des mesures a posteriori  $p(\underline{\theta}_p/\underline{Y})$  maximale.

$p(\underline{\theta}_p/\underline{Y})$  est appelée fonction de vraisemblance. La solution  $\underline{\theta}$  est celle qui donne la probabilité maximale de trouver avec le modèle la valeur des mesures expérimentales obtenues sur le système.

Supposons que le modèle soit de la forme:

$$\underline{Y} = X \underline{\theta} + \underline{b}$$

où X est la matrice des valeurs déterministes et  $\underline{b}$  le bruit supposé gaussien<sup>(1)</sup>, centré. La densité de probabilité de  $\underline{b}$  est par définition:

$$p(\underline{b}) = \frac{1}{2\pi^{\frac{N+1}{2}} |R|^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2} \underline{b}^T \cdot R^{-1} \cdot \underline{b}\right)$$

avec :

$$\begin{aligned} E.M.(\underline{b}) &= \underline{0} \\ E.M.(\underline{b} \cdot \underline{b}^T) &= R \end{aligned}$$

N; nombre de mesures (dimension de  $\underline{Y}$  et de  $\underline{b}$ )

alors:

$$p(\underline{Y}/\underline{\theta}_p) = \frac{1}{2\pi^{\frac{N+1}{2}} |R|^{\frac{1}{2}}} \cdot \exp\left[-\frac{1}{2} (\underline{Y} - X \underline{\theta}_p)^T R^{-1} (\underline{Y} - X \underline{\theta}_p)\right]$$

(1) Etant donné que la plupart des perturbations peuvent être considérées comme résultant d'un grand nombre de causes indépendantes, en vertu du théorème central limite, elles sont gaussiennes. Toutes leurs statistiques peuvent être déterminées à partir de la valeur moyenne et de la fonction d'autocorrélation.

dont le maximum est donné par:

$$-\frac{\partial}{\partial \underline{\theta}_p} \left[ (\underline{Y} - \underline{X} \underline{\theta}_p)^T R^{-1} (\underline{Y} - \underline{X} \underline{\theta}_p) \right] = 0$$

d'où:

$$\hat{\underline{\theta}} = (\underline{X}^T R^{-1} \underline{X})^{-1} \underline{X}^T R^{-1} \underline{Y}$$

Cet estimateur est appelé estimateur de **MARKOV**. Il exige la connaissance de R, matrice de covariance du bruit. Si le bruit est blanc<sup>(2)</sup>, il devient :

$$\hat{\underline{\theta}} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{Y}$$

qui est l'estimateur des moindres carrés.

On voit donc que des hypothèses de plus en plus restrictives sur le système permettent de déduire les divers estimateurs les uns des autres.

On doit remarquer que là encore, l'influence de l'entrée est très importante: en effet, celle-ci doit varier suffisamment autour d'une valeur moyenne finie pour pouvoir exciter tous les paramètres du système. Nous verrons plus loin quel type de signal nous devons adopter pour exciter le système à identifier.

---

(2) Un bruit est dit blanc, par définition, si son spectre est constant. Un tel signal n'a pas d'existence physique car il correspond à une énergie infinie, dans le cas de signaux continus. Par extension, un signal échantillonné est dit pseudo-blanc si son spectre est constant et limité. Dans ce cas, son énergie est finie et il est facile de construire un signal présentant ces caractéristiques. On montre qu'un bruit pseudo-blanc implique l'indépendance des échantillons entre eux. Cette possibilité de fabriquer un bruit pseudo-blanc (par tirage au sort, de façon indépendante, d'une suite de chiffres), donne à ce signal un sens concret que n'a pas son homologue continu, le bruit blanc, qui n'est qu'un être mathématique irréalisable.

CHAPITRE II

E T U D E   D U   M O D E L E

II-1   ANALYSE DU MODELE  
=====

Parmi les modèles les plus courants, celui que nous nous proposons d'utiliser est de la forme :

$$y(k+n) + a_1 y(k+n-1) + \dots + a_n y(k) = b_1 u(k+p) + \dots + b_p u(k) \quad (\text{II-1})$$

Sa nature discrète présente un avantage sur le modèle continu du type :

$$y(t) + a_1 y^{(1)}(t) + \dots + a_n y^{(n)}(t) = b_1 u(t) + \dots + b_p u^{(p)}(t) \quad (\text{II-2})$$

En effet, les mesures effectuées sur le procédé sont fournies par un ordinateur couplé à ce procédé. L'information ainsi obtenue est discrète.

D'autre part, le calculateur intervient dans l'identification elle-même; le traitement de l'information est donc effectué de façon discrète.

Le problème majeur soulevé par un modèle du type (II-2) est la détermination de l'ordre de ce modèle. L'ordre du système à identifier est inconnu au départ. Celui-ci peut être dans certains cas à paramètres répartis, donc théoriquement d'ordre infini, que l'on assimile à un système linéaire d'ordre fini. La difficulté principale réside dans le choix de l'ordre à attribuer au modèle pour que celui-ci copie correctement le fonctionnement du système.

Essayons d'examiner maintenant de près la structure du modèle : soit le système schématisé par la figure (II-1)

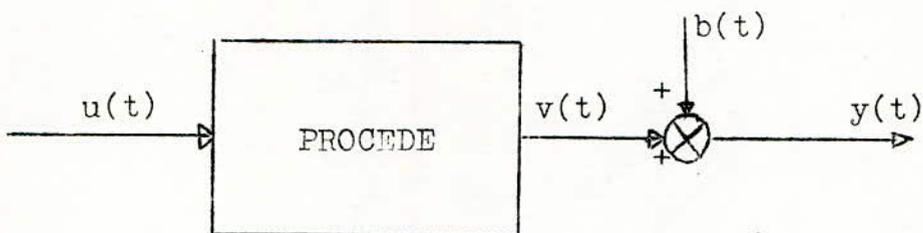


figure II-1

$u(t)$  : entrée du procédé

$v(t)$  : sortie déterministe du procédé

$b(t)$  : bruit additif en sortie, dû au bruit de mesure et au bruit de fonctionnement; son additivité en sortie est possible grâce à l'hypothèse de linéarité du système

$y(t)$  : sortie mesurée du procédé.

On supposera que les mesures de l'entrée ne sont entachées d'aucun bruit. Le calculateur fournissant ces mesures est en ligne sur le système et les erreurs de quantification sont négligeables.

On supposera de plus, que le procédé est stationnaire.

L'équation aux différences décrivant le fonctionnement du système, sous sa forme la plus générale, est:

$$v(t) + a_1 v(t-T) + \dots + a_n v(t-nT) = b_0 u(t) + b_1 u(t-T) + \dots + b_p u(t-pT)$$

où  $n$  est l'ordre et  $T$  la période d'échantillonnage du système. pour simplicité d'écriture, nous supposons dans tout ce qui suit  $T$  unitaire :  $T=1$

$$\Rightarrow v(t) + a_1 v(t-1) + \dots + a_n v(t-n) = b_0 u(t) + b_1 u(t-1) + \dots + b_p u(t-p)$$

avec  $p \leq n$  pour condition de réalisabilité physique du système, car la sortie  $v(t-n)$  ne peut dépendre des valeurs de l'entrée aux instants ultérieurs.

Avec l'hypothèse de bruit ramené en sortie, nous avons:

$$y(t) = v(t) + b(t)$$

L'équation du système devient:

$$y(t) + \sum_{i=1}^n a_i y(t-i) = \sum_{i=0}^p b_i u(t-i) + b(t) + \sum_{i=1}^n a_i b(t-i) \quad (II-3)$$

En introduisant l'opérateur de retard  $Z^{-i}$  défini par:

$$x(t-i) = Z^{-i} x(t) \quad , \text{ il vient :}$$

$$y(t) + \sum_{i=1}^n a_i Z^{-i} y(t) = \sum_{i=0}^p b_i Z^{-i} u(t) + b(t) + \sum_{i=1}^n a_i Z^{-i} b(t)$$

ou encore :

$$A(Z^{-1}) y(t) = B(Z^{-1}) u(t) + A(Z^{-1}) b(t) \quad (II-4)$$

où :

$$A(Z^{-1}) = 1 + a_1 Z^{-1} + \dots + a_n Z^{-n}$$

$$B(Z^{-1}) = b_0 + b_1 Z^{-1} + \dots + b_p Z^{-p}$$

Rappelons que le système échantillonné est stable si les racines du polynôme  $A(Z^{-1})$  appartiennent toutes au domaine intérieur au cercle de rayon unité dans le plan des  $Z$ .

De plus,  $A(Z^{-1})$  et  $B(Z^{-1})$  sont supposés n'avoir aucun facteur commun: ceci correspond à la condition d'observabilité<sup>(1)</sup> et de commandabilité<sup>(2)</sup> du système.

Un système représenté par l'équation vectorielle d'état:

$$\underline{x}(k+1) = f [\underline{x}(k), \underline{u}(k), k]$$

$$\underline{y}(k) = g [\underline{x}(k), k]$$

est dit :

(1) observable si  $\forall k_0, k_1 (k_1 \geq k_0)$ , tel que l'observation de  $\underline{y}(k)$  et la connaissance de  $\underline{u}(k)$  sur l'intervalle temporel  $[k_0, k_1]$ , permettent de déterminer l'état initial  $\underline{x}(k_0)$ .

(2) commandable si, étant donné  $k_0$  quelconque et deux états quelconques  $\underline{x}_0$  et  $\underline{x}_1$ , on peut trouver un instant  $k_1 (k_1 \geq k_0)$ , et une commande admissible  $\underline{u}(k)$  sur l'intervalle  $[k_0, k_1]$ , transférant le système de l'état  $\underline{x}_0$  à l'instant  $k_0$ , à l'état  $\underline{x}_1$  à l'instant  $k_1$ .

II-2 LES RESIDUS

Le modèle peut être mis sous la forme suivante:

$$A(Z^{-1}) y(t) = B(Z^{-1}) u(t) + e(t) \quad (II-5)$$

avec : 
$$e(t) = A(Z^{-1}) b(t) \quad (II-6)$$

$e(t)$  représentant un bruit généralisé, appelé résidu.  
Que représentent ces résidus? On a :

$$e(t) = y(t) + a_1 y(t-1) + \dots + a_n y(t-n) - b_0 u(t) - \dots - b_p u(t-p) \quad (II-7)$$

Il est possible de prédire la sortie du procédé à l'aide du modèle, des mesures précédentes et de la consigne. Nous noterons cette prédiction à l'instant  $t$  :  $\hat{y}(t/t-1)$

$$\hat{y}(t/t-1) = - a_1 y(t-1) - \dots - a_n y(t-n) + b_0 u(t) + \dots + b_p u(t-p) \quad (II-8)$$

REMARQUE : le concept de sortie prédite par le modèle diffère de celui de sortie du modèle :

-La prédiction à l'instant  $t$  est déduite de la mesure des entrées et des sorties aux  $N$  instants précédents,

-La sortie du modèle  $y_m$  est calculée à partir de l'équation récurrente :

$$A(Z^{-1}) y_m(t) = B(Z^{-1}) u(t)$$

depuis l'instant zéro, avec  $n$  conditions initiales.

A partir des équations (II-7) et (II-8), on tire :

$$e(t) = y(t) - \hat{y}(t/t-1) \quad (II-9)$$

Ce résidu représente l'erreur de prédiction du modèle. Si les paramètres du modèle et du système sont identiques, cette erreur est uniquement fonction du bruit sur le système :

$$e(t) = A(Z^{-1}) b(t)$$

Dans le cas contraire, elle est aussi fonction de l'imprécision des paramètres. Elle peut dépendre encore du mauvais choix de l'ordre du modèle. Elle aura dans ce cas peu de lien avec le bruit  $b(t)$ .

### II-3 LE RETARD PUR =====

Dans la pratique, on rencontre souvent des systèmes comprenant un grand nombre de petites constantes de temps que l'on assimile à un retard pur. Parfois il s'agit même d'un véritable retard existant dans le système.

Le modèle des procédés industriels nécessite souvent l'introduction de ce retard ou temps mort. Celui-ci peut avoir une cause physique: transport de matière par exemple. Sur le schéma (figure II-2), la première constante de temps d'un système du second ordre est très faible et pourra facilement être approchée par un retard pur  $R_e$  :

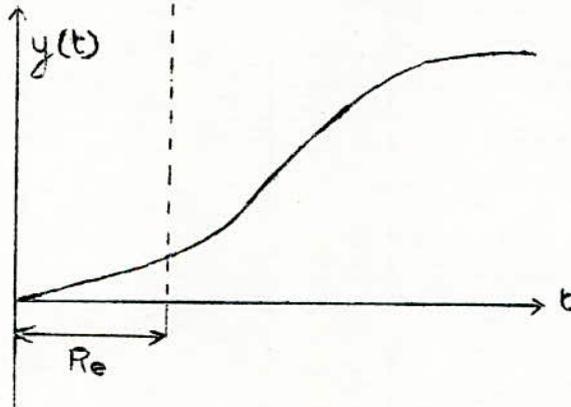


figure II-2

Ce retard pur sera compté en nombre entier de périodes d'échantillonnage :

$$R_e = R.T = R \quad ( T \text{ unitaire } )$$

L'équation (II-3) devient :

$$y(t) + \sum_{i=1}^n a_i y(t-i) = \sum_{i=0}^p b_i u(t-i-R) + b(t) + \sum_{i=1}^n a_i b(t-i)$$

et l'équation (II-5) :

$$A(Z^{-1}) y(t) = Z^{-R} B(Z^{-1}) u(t) + e(t)$$

Le retard se traduit par une simple translation de l'entrée. Nous verrons plus loin comment déterminer pratiquement R, qui n'est pas considéré comme un paramètre à identifier, mais comme une donnée a priori. Pour simplifier les écritures, nous supposons dans tout ce qui suit  $R = 0$ .

II-4 REPRESENTATION DE LA FONCTION DE TRANSFERT ECHANTILLONNEE ;  
 =====

GAIN STATIQUE  
 =====

Pour un modèle représenté sous la forme d'une équation aux différences :

$$y(t) + \sum_{i=1}^n a_i y(t-i) = \sum_{i=0}^p b_i u(t-i) \quad (II-10)$$

la transformée en Z d'un signal  $y(t)$  est :

$$Z(y(t)) = \sum_{n=0}^{\infty} y(n) Z^{-n} = Y(Z)$$

$$y(t) = 0 \quad \text{pour } t < 0$$

$$\begin{aligned} Z(y(t-k)) &= \sum_{n=0}^{\infty} y(n-k) Z^{-n} = Z^{-k} \sum_{n=0}^{\infty} y(n-k) Z^{-n+k} \\ &= Z^{-k} \sum_{m=-k}^{\infty} y(m) Z^{-m} = Z^{-k} \sum_{m=0}^{\infty} y(m) Z^{-m} \\ &= Z^{-k} Y(Z) \end{aligned}$$

Donc la transformée en Z de l'équation récurrente (II-10) est:

$$Y(Z) + \sum_{i=1}^n a_i Z^{-i} Y(Z) = \sum_{i=0}^p b_i Z^{-i} U(Z)$$

La fonction de transfert échantillonnée du système (définie pour des signaux nuls si  $t < 0$ ) est :

$$F(Z) = \frac{Y(Z)}{U(Z)} = \frac{\sum_{i=0}^p b_i Z^{-i}}{1 + \sum_{i=1}^n a_i Z^{-i}}$$

Le gain statique du système est alors :

$$G = \frac{\sum_{i=0}^p b_i}{1 + \sum_{i=1}^n a_i}$$

On montre que lors de l'échantillonnage, c'est-à-dire du passage du domaine continu (plan  $p$ ) au domaine discret (plan  $Z$ ), les paramètres sont plus précis dans le plan  $Z$  que dans le plan  $p$ . Ainsi, pour obtenir une sortie équivalente avec deux fonctions de transfert, par exemple :

$$\frac{Z}{Z - a} \quad \text{et} \quad \frac{1}{p - \text{Log } a/T} = \frac{K}{1 + \tau p}$$

le pôle  $Z$  doit être déterminé avec plus de précision que le pôle en  $p$ .

## II-5 PROBLEMES D'ECHANTILLONNAGE

Rappelons le théorème de SHANNON :

Pour pouvoir reconstituer un signal continu à partir d'un train d'échantillons de période  $T_e$ , il faut que la pulsation d'échantillonnage  $\omega_e = 2\pi/T_e$  soit au moins deux fois plus grande que la plus grande des pulsations contenues dans le spectre du signal que l'on échantillonne.

En pratique, il suffit de prendre  $\omega_e$  largement supérieure à la pulsation de coupure du système, car le spectre du signal n'est pas limité; mais il ne faut pas oublier qu'un échantillonnage à période longue permet de traiter séquentiellement un grand nombre de signaux par un seul calculateur. Le choix de cette période résulte d'un compromis à faire entre stabilité et précision du système d'une part, et coût du calculateur (lié à sa rapidité et sa capacité), d'autre part. Connaissant les

caractéristiques dynamiques du processus à régler, il est possible de calculer la période maximale permise .

Les problèmes d'échantillonnage se posent surtout pour l'identification en temps réel et la commande du processus :

- C'est dans un intervalle de temps inférieur ou égal à la période d'échantillonnage que s'effectue l'identification dynamique du procédé.

- C'est dans cet intervalle de temps qu'est élaboré un signal d'action (fonction notamment du dernier échantillon), dans le cas de la commande.

## II-6 STABILITE =====

Considérons le système décrit par l'équation :

$$a_0 y(kT) + a_1 y[(k+1)T] + \dots + a_n y[(k+n)T] = b_0 u(kT) + b_1 u[(k+1)T] + \dots + b_p u[(k+p)T]$$

soit : 
$$\sum_{i=0}^n a_i y[(k+i)T] = \sum_{i=0}^p b_i u[(k+i)T] \quad \text{avec } p \leq n$$

Prenons la transformée en Z des deux membres et supposons les conditions initiales nulles :

$$Y(Z) \sum_{i=0}^n a_i Z^{+i} = U(Z) \sum_{i=0}^p b_i Z^i$$

on a donc :

$$Y(Z) = U(Z) \frac{\sum_{i=0}^p b_i Z^i}{\sum_{i=0}^n a_i Z^i} = U(Z) \frac{B(Z)}{A(Z)}$$

$B(Z)/A(Z) = H(Z)$  représente la transmittance échantillonnée qui lie  $U(Z)$  et  $Y(Z)$ .  $Y(Z)$  peut se mettre sous la forme :

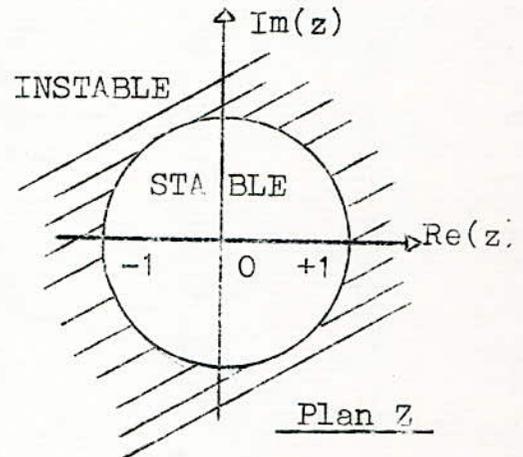
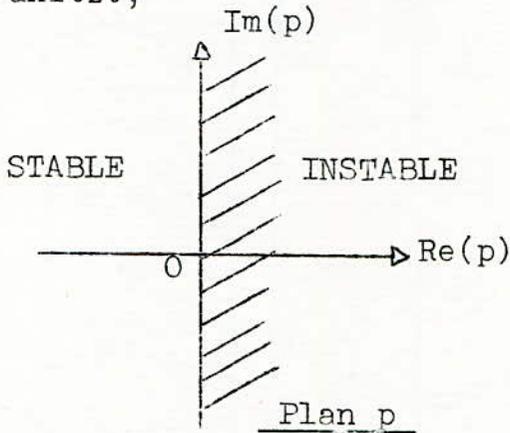
$$Y(Z) = \frac{R_1}{1 - \alpha_1 Z^{-1}} + \frac{R_2}{1 - \alpha_2 Z^{-1}} + \dots$$

où les  $R_i$  sont les résidus de la fonction relatifs aux pôles  $\alpha_i$  la forme  $y(t)$  aux instants  $nT$  sera, d'après le théorème d'inversion de la transformée en Z :

$$y(nT) = \sum_{i=1}^n R_i \alpha_i^{nT}$$

Ce signal tendra vers l'état d'équilibre zéro si tous les  $\alpha_i$  ont leur module inférieur à 1.

Le système est donc stable si toutes les racines du polynôme  $A(Z)$  (ou de l'équation caractéristique  $A(Z)=0$ ), c'est-à-dire les pôles de la transmittance, sont à l'intérieur du cercle unité.;



REMARQUE : ceci est l'expression dans le plan de la variable Z des conditions de stabilité d'un système linéaire dans le plan de la variable p

En effet:  $R_e(p_i) < 0$  ;  $(z_i = e^{Tp_i}) \Rightarrow |z_i| < 1$

L'étude de la stabilité d'un système échantillonné linéaire revient donc à l'étude du module des pôles de sa transmittance  $H(Z) = B(Z)/A(Z)$ , c'est-à-dire des racines de l'équation caractéristique  $A(Z) = 0$ .

CHAPITRE III

M E T H O D E D E S M O I N D R E S C A R R E S

Avant de passer à l'étude de la méthode, il est important de faire la distinction entre l'identification en temps réel et l'identification en temps différé. Les précisions que nous apportons ci-dessous seront valables pour les autres méthodes statistiques d'identification que nous étudierons dans les chapitres suivants.

L'identification d'un processus industriel en temps différé ne nécessite pas la présence d'un ordinateur en ligne (on-liné) avec ce processus. En effet, les données sont fournies par des mesures prises sur le procédé et l'identification est faite par la suite, d'une façon indépendante.

De nos jours, on s'intéresse de plus en plus à des méthodes qui identifient les procédés en temps réel, c'est-à-dire que le modèle est mis à jour au fur et à mesure que les données sont acquises. Ces techniques sont indispensables, surtout lorsque l'on a affaire à des systèmes dont les paramètres varient effectivement dans le temps; de plus, elles présentent un grand intérêt lorsque le système étudié est non linéaire, mais traité comme un système linéaire autour d'un point de fonctionnement; les valeurs des paramètres dépendent alors du point de fonctionnement. Elles peuvent de la même façon dépendre des perturbations. D'autre part, l'étude de l'évolution des paramètres d'un système au cours du temps, même s'il est linéaire, est particulièrement riche en renseignements car elle permet l'étude du vieillissement de l'installation, ce qui peut permettre une meilleure gestion de l'entreprise.

# A / OPTIQUE TEMPS DIFFERE

Le modèle retenu a pour équation (cf II-5) :

$$A(Z^{-1}) y(t) = B(Z^{-1}) u(t) + e(t)$$

ou bien :  $y(t) + a_1 y(t-1) + \dots + a_n y(t-n) = b_0 u(t) + \dots + b_p u(t-p) + e(t)$   
 où  $e(t)$  est l'erreur de prédiction du modèle.

### III-1 LE CRITERE A MINIMISER ET LA SOLUTION

=====

Avec N mesures de l'entrée et de la sortie, entre les instants 1 et N, on peut écrire :

$$\begin{aligned} y(N) &= -a_1 y(N-1) - \dots - a_n y(N-n) + b_0 u(N) + \dots + b_p u(N-p) + e(N) \\ y(N-1) &= -a_1 y(N-2) - \dots - a_n y(N-n-1) + b_0 u(N-1) + \dots + b_p u(N-p-1) + e(N-1) \\ &\vdots \\ y(n+1) &= -a_1 y(n) - \dots - a_n y(1) + b_0 u(n+1) + \dots + b_p u(n-p+1) + e(n+1) \end{aligned}$$

Pour que le système soit physiquement réalisable, on doit avoir :

$$p \leq n$$

Ces N-n équations peuvent s'écrire sous la forme matricielle

$$\begin{bmatrix} y(N) \\ y(N-1) \\ \vdots \\ y(n+1) \end{bmatrix} = \begin{bmatrix} -y(N-1) \dots - y(N-n) & u(N) & \dots & u(N-p) \\ -y(N-2) \dots - y(N-n-1) & u(N-1) & \dots & u(N-p-1) \\ \vdots & \vdots & \ddots & \vdots \\ -y(n) \dots - y(1) & u(n+1) & \dots & u(n+1-p) \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \\ b_0 \\ \vdots \\ b_p \end{bmatrix} + \begin{bmatrix} e(N) \\ e(N-1) \\ \vdots \\ e(n+1) \end{bmatrix} \quad (\text{III-1})$$

On voit que pour un tel modèle, n mesures de la sortie et p+1 mesures de l'entrée interviennent sous forme de condi-

tions initiales et qu'avec N mesures entrée-sortie on obtient N-n équations. Il n'est donc pas nécessaire de poser les conditions initiales nulles pour cette méthode.

L'équation matricielle (III-1) peut encore s'écrire:

$$\underline{Y} = X \underline{\theta} + \underline{E} \quad (\text{III-2})$$

où  $\underline{\theta}$  est le vecteur des paramètres à estimer.

Le critère à minimiser est le carré de l'erreur de prédiction :

$$C = \underline{E}^T \underline{E} = \sum_{t=n+1}^N e^2(t)$$

soit:

$$\begin{aligned} C &= \underline{E}^T \underline{E} = (\underline{Y} - X \underline{\theta})^T (\underline{Y} - X \underline{\theta}) \\ &= \underline{Y}^T \underline{Y} - \underline{\theta}^T X^T \underline{Y} - \underline{Y}^T X \underline{\theta} + \underline{\theta}^T X^T X \underline{\theta} \\ &= \underline{Y}^T \underline{Y} - 2 \underline{Y}^T X \underline{\theta} + \underline{\theta}^T X^T X \underline{\theta} \end{aligned}$$

L'estimateur  $\hat{\underline{\theta}}$  qui minimise ce critère découle de la condition nécessaire suivante :

$$\frac{\partial C}{\partial \underline{\theta}} = -2 \underline{Y}^T X + 2 \underline{\theta}^T X^T X = 0$$

$$\implies \boxed{\hat{\underline{\theta}} = (X^T X)^{-1} X^T \underline{Y}} \quad (\text{III-3})$$

La matrice des dérivées secondes est :

$$C_{\theta\theta} = \left\{ \frac{\partial^2 C}{\partial \theta_i \partial \theta_j} \right\} = 2 X^T X$$

c'est une matrice définie positive, symétrique. La solution  $\hat{\underline{\theta}}$  est donc bien un minimum.

L'avantage qu'offre cette méthode sur le plan du calcul est que l'estimateur des paramètres est explicite. Son calcul suppose néanmoins que la matrice  $X^T X$  soit régulière.

Ce qui précède appelle quelques remarques :

- a) Le modèle que nous avons choisi est linéaire par rapport aux

paramètres; c'est ce qui importe dans le calcul de l'estimateur. Il pourrait représenter un système non linéaire par rapport aux entrées-sorties, par exemple:

$$y(t) = a y(t-1) + b u^2(t) + e(t)$$

Dans ce cas, pour le calcul de la solution (III-3), il suffit de poser:

$$X = \begin{bmatrix} y(N-1) \dots u^2(N) \\ \vdots \\ y(1) \dots u^2(1) \end{bmatrix} \quad \underline{\theta} = \begin{bmatrix} a \\ b \end{bmatrix}$$

- b) si le système n'était pas bruité, il aurait suffi de résoudre un système à  $n+p+1$  équations:

$$\begin{aligned} y(n+1) + a_1 y(n) + \dots + a_n y(1) &= b_0 u(n+1) + \dots + b_p u(n-p+1) \\ \vdots & \\ y(2n+1+p) + a_1 y(2n+p) + \dots + a_n y(n+1+p) &= b_0 u(2n+1+p) + \dots + b_p u(2n) \end{aligned}$$

Du fait de la présence du bruit, on est amené à résoudre  $N-n$  équations à  $n+p+1$  inconnues, ceci afin d'obtenir un effet de moyennage du bruit.

- c) Cette méthode ne nécessite pas une entrée  $u(t)$  particulière mais on a intérêt à la choisir de façon à exciter le plus possible le système sans perturber son fonctionnement.

### III-2 CARACTERISTIQUES DE L'ESTIMATEUR

=====

Un estimateur est dit non biaisé si:

$$E.M \{ \hat{\theta} \} = \underline{\theta}$$

Il converge alors en moyenne vers la vraie valeur des paramètres; cette qualité est évidemment indispensable pour une bonne identification. Voyons si elle est vérifiée pour l'estimateur des moindres carrés:

$$\begin{aligned}
 E.M\{\hat{\underline{\theta}}\} &= E.M \left\{ (X^T X)^{-1} X^T \underline{Y} \right\} \\
 &= E.M \left\{ (X^T X)^{-1} X^T (X \underline{\theta} + \underline{E}) \right\} \\
 &= E.M \left\{ (X^T X)^{-1} X^T X \underline{\theta} + (X^T X)^{-1} X^T \underline{E} \right\} \\
 &= \underline{\theta} + E.M \left\{ (X^T X)^{-1} X^T \underline{E} \right\}
 \end{aligned}$$

Ainsi, l'estimateur est non biaisé si seulement:

$$E.M \left\{ \underline{E} \right\} = 0$$

c'est-à-dire si  $\underline{E}$  est non corrélé avec  $X$  et est centré; alors:

$$E.M \left\{ \hat{\underline{\theta}} \right\} = \underline{\theta}$$

La corrélation des paramètres est donnée par la matrice de covariance de l'estimateur  $\Gamma_{\underline{\theta}\underline{\theta}}$ . Les éléments  $\gamma_{\theta_i \theta_j}$  de cette matrice indiquent la corrélation entre paramètres  $\theta_i$  et  $\theta_j$  et les éléments diagonaux  $\gamma_{\theta_i \theta_i}$  la variance de chaque paramètre qui est une mesure de la précision de l'estimation.

$$\begin{aligned}
 \Gamma_{\underline{\theta}\underline{\theta}} &= E.M \left\{ (\hat{\underline{\theta}} - \underline{\theta}) \cdot (\hat{\underline{\theta}} - \underline{\theta})^T \right\} \\
 &= E.M \left\{ (X^T X)^{-1} X^T \underline{E} \underline{E}^T X (X^T X)^{-1} \right\}
 \end{aligned}$$

A condition que le bruit soit blanc, de variance  $\sigma^2$  c'est-à-dire si:

$$E.M \left\{ \underline{E} \underline{E}^T \right\} = \sigma^2 I$$

on démontre que:

$$\Gamma_{\underline{\theta}\underline{\theta}} = \sigma^2 (X^T X)^{-1} \tag{III-4}$$

Si la variance des résidus n'est pas connue, un estimateur non biaisé de cette variance est donné par:

$$\hat{\sigma}^2 = \frac{1}{M - m} \hat{\underline{E}}^T \hat{\underline{E}} = \frac{1}{M - m} (\underline{Y} - X \hat{\underline{\theta}})^T (\underline{Y} - X \hat{\underline{\theta}}) \tag{III-5}$$

où  $M = N - n$  et  $m = n + p + 1$ , sont les dimensions de la matrice  $X$  et  $\hat{\underline{E}}$  l'estimation des résidus par le modèle.

Il est essentiel de remarquer que pour l'étude des

propriétés statistiques de l'estimateur, nous avons exigé deux conditions :

-E centré

-E non corrélé avec X

On peut toujours vérifier la première hypothèse en considérant un paramètre supplémentaire dans l'équation du modèle, pour centrer les résidus. Par contre, d'après (II-6) nous savons que:

$$e(t) = A(Z^{-1}).b(t)$$

Les résidus sont donc toujours corrélés par l'intermédiaire de  $A(Z^{-1})$ . Il intervient dans l'expression de  $e(t)$  des termes en  $b(t-1), b(t-2), \dots, b(t-n)$ . et sur la ligne correspondante de la la matrice X, on trouve les termes  $y(t-1), y(t-2), \dots, y(t-n)$ , qui sont fonction de  $b(t+1), \dots, b(t-n)$ .

Donc avec le modèle choisi, la corrélation est inévitable. L'estimateur  $\hat{\theta}$  obtenu par la méthode des moindres carrés est toujours biaisé. La non corrélation ne peut être respectée que si le résidu  $e(t)$  est blanc. L'hypothèse de bruit en sortie du système  $b(t)$  blanc ne s'avère donc pas suffisante.

Soulignons encore que ce problème de biais de l'estimateur est un problème dû au modèle choisi plus qu'à la méthode elle-même.

Dans cette optique temps différé, la solution du problème des moindres carrés présente au moins l'avantage énorme d'être explicite et très facile à mettre en oeuvre.

La méthode des moindres carrés en temps réel se déduit facilement de l'optique temps différé.

Soit N le nombre de mesures effectuées jusqu'à l'instant  $t=NT=N$  (T unitaire). Nous avons vu que l'estimateur des moindres carrés a pour expression (III-3) :

$$\hat{\underline{\theta}}_N = (\underline{X}_N^T \underline{X}_N)^{-1} \underline{X}_N^T \underline{Y}_N$$

Si on acquiert une mesure supplémentaire, cet estimateur devient :

$$\hat{\underline{\theta}}_{N+1} = (\underline{X}_{N+1}^T \underline{X}_{N+1})^{-1} \underline{X}_{N+1}^T \underline{Y}_{N+1}$$

avec :

$$\underline{X}_{N+1} = \begin{bmatrix} \underline{X}_N \\ x(N+1) \end{bmatrix} \quad \underline{Y}_{N+1} = \begin{bmatrix} \underline{Y}_N \\ y(N+1) \end{bmatrix}$$

$$x(N+1) = \begin{bmatrix} -y(N) \dots -y(N-n+1) & u(N+1) \dots u(N+1-p) \end{bmatrix}$$

$$\underline{X}_{N+1}^T \underline{X}_{N+1} = \begin{bmatrix} \underline{X}_N^T & x^T(N+1) \end{bmatrix} \begin{bmatrix} \underline{X}_N \\ x(N+1) \end{bmatrix}$$

$$= \underline{X}_N^T \underline{X}_N + x^T(N+1) x(N+1)$$

$$\underline{X}_{N+1}^T \underline{Y}_{N+1} = \underline{X}_N^T \underline{Y}_N + x^T(N+1) y(N+1)$$

$$\implies \hat{\underline{\theta}}_{N+1} = \left[ \underline{X}_N^T \underline{X}_N + x^T(N+1) x(N+1) \right]^{-1} \left[ \underline{X}_N^T \underline{Y}_N + x^T(N+1) y(N+1) \right]$$

En utilisant la formule de calcul matriciel suivante:

$$(A + BCD)^{-1} = A^{-1} - A^{-1} B (C^{-1} + D A^{-1} B)^{-1} D A^{-1}$$

En posant:  $A = X_N^T X_N$ ,  $B = x^T(N+1)$ ,  $C = I$ ,  $D = x(N+1)$   
on a, en omettant l'indice N ou N+1, pour commodité de calcul:

$$\begin{aligned} \left[ X_N^T X_N + x^T(N+1) x(N+1) \right]^{-1} &= (X^T X)^{-1} - (X^T X)^{-1} x^T \left[ 1 + x(X^T X)^{-1} x^T \right]^{-1} x (X^T X)^{-1} \\ \text{d'où : } \hat{\underline{\theta}}_{N+1} &= \hat{\underline{\theta}}_N - \left[ (X^T X)^{-1} x^T (1 + x(X^T X)^{-1} x^T)^{-1} x \hat{\underline{\theta}}_N + (X^T X)^{-1} x^T y - \right. \\ &\quad \left. - (X^T X)^{-1} x^T (1 + x(X^T X)^{-1} x^T)^{-1} x (X^T X)^{-1} x^T y \right] \\ &= \hat{\underline{\theta}}_N - (X^T X)^{-1} x^T \left[ 1 + x(X^T X)^{-1} x^T \right]^{-1} \left[ x \hat{\underline{\theta}}_N + x(X^T X)^{-1} x^T y \right] + \\ &\quad + (X^T X)^{-1} x^T y \end{aligned}$$

la quantité  $x(X^T X)^{-1} x^T$  est un scalaire; nous avons finalement:

$$\hat{\underline{\theta}}_{N+1} = \hat{\underline{\theta}}_N + \frac{(X_N^T X_N)^{-1} x^T(N+1)}{1 + x(N+1)(X_N^T X_N)^{-1} x^T(N+1)} \left[ y(N+1) - x(N+1) \hat{\underline{\theta}}_N \right]$$

Posons  $P_N = (X_N^T X_N)^{-1}$  que l'on appellera matrice de gain. En remarquant que cette matrice peut être elle aussi calculée récursivement, on a:

$$P_{N+1} = P_N - \underline{K}_{N+1} x(N+1) P_N$$

où  $\underline{K}_{N+1}$  est un gain dont l'expression est:

$$\underline{K}_{N+1} = P_N x^T(N+1) \left[ 1 + x(N+1) P_N x^T(N+1) \right]^{-1}$$

L'algorithm final est alors:

$$\begin{aligned}
 \hat{\underline{\theta}}_{N+1} &= \hat{\underline{\theta}}_N + \underline{K}_{N+1} ( y(N+1) - x(N+1) \hat{\underline{\theta}}_N ) \\
 \underline{K}_{N+1} &= P_N x^T(N+1) \left[ 1 + x(N+1) P_N x^T(N+1) \right]^{-1} \\
 P_{N+1} &= P_N - \underline{K}_{N+1} x(N+1) P_N
 \end{aligned}
 \tag{III-6}$$

Les matrices  $\hat{\underline{\theta}}, \underline{K}$  et  $P$  peuvent être mises à jour à chaque instant d'échantillonnage grâce à l'acquisition de  $y(N+1)$  et  $u(N+1)$ . IL faut noter que le calcul ne nécessite aucune inversion de matrice puisque la quantité  $x(N+1) P_N x^T(N+1)$  est un scalaire

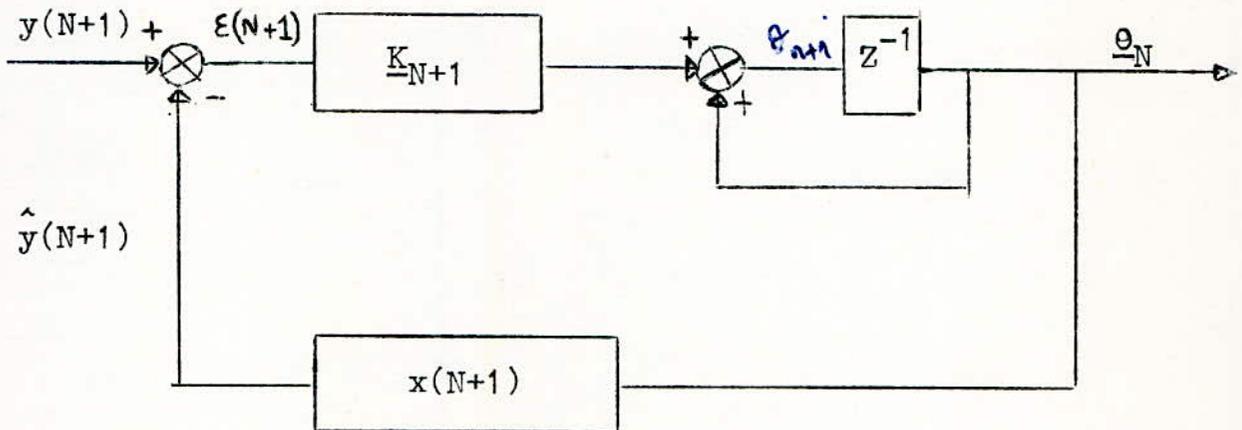
Pour initialiser la récurrence (III-6),  $n+p+1$  échantillons seraient nécessaires, en principe, pour construire la matrice  $P_0$  de départ. En fait, on montre que la valeur initiale  $\hat{\underline{\theta}}_0$  des paramètres a peu d'importance et qu'une initialisation de  $P_0$  à  $1/\epsilon$  ( $\epsilon \rightarrow 0$ ), c'est-à-dire une matrice dont tous les éléments sont nuls sauf la diagonale qui est infinie, permet la convergence de l'algorithme.

--- Interprétation physique de l'algorithme ---

Le nouveau vecteur  $\hat{\underline{\theta}}_{N+1}$  est égal à l'ancien  $\hat{\underline{\theta}}_N$  à un terme correctif près. Ce terme est le produit par un gain  $\underline{K}_{N+1}$  de l'expression:

$$y(N+1) - x(N+1) \hat{\underline{\theta}}_N = y(N+1) - \hat{y}(N+1/N)$$

Cette expression représente l'erreur de prédiction de l'estimateur  $\hat{\theta}_N$ , ce à quoi on s'attendait, puisque l'estimateur des moindres carrés minimise l'erreur de prédiction. L'algorithme peut être représenté par le système à contre-réaction suivant :



La sortie de ce système,  $\hat{\theta}_N$ , est corrigée avec le gain  $K_{N+1}$  tant que l'erreur  $\epsilon_{N+1}$  entre sortie mesurée et sortie prédite n'est pas nulle.

CHAPITRE IV

M E T H O D E  
D E L A  
M A T R I C E I N S T R U M E N T A L E

Rappelons l'équation du système :

$$y(t) + \sum_{i=1}^n a_i y(t-i) = \sum_{i=0}^p b_i u(t-i) + e(t)$$

soit :  $A(z^{-1}) y(t) = B(z^{-1}) u(t) + e(t)$

Avec N mesures de l'entrée et de la sortie , on peut écrire:

$$\underline{Y} = X \underline{\theta} + \underline{E}$$

$\underline{Y}$  est le vecteur des sorties, X la matrice formée des données entrée-sortie,  $\underline{E}$  le vecteur des résidus ou bruit généralisé et  $\underline{\theta}$  le vecteur des paramètres à estimer.

Les dimensions de ces tableaux sont:

$$\dim(\underline{Y}) = M \times 1$$

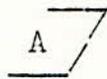
$$\dim(X) = M \times m$$

$$\dim(\underline{\theta}) = m \times 1$$

$$\dim(\underline{E}) = M \times 1$$

avec  $\begin{cases} m = n+p+1 \\ M = N-n \end{cases}$

Pour l'étude de cette méthode, nous allons procéder de la même manière que pour celle des moindres carrés . De l'étude en temps différé nous déduirons l'algorithme d'identification en temps réel.



IV-1 PRINCIPE DE LA METHODE  
=====

Soit une matrice Z de dimension Mxm ; on peut écrire:

$$Z^T \underline{Y} = Z^T X \underline{\theta} + Z^T \underline{E}$$

Si Z est telle que l'inverse de  $(Z^T X)$  existe, alors:

$$\underline{\theta} = (Z^T X)^{-1} Z^T \underline{Y} - (Z^T X)^{-1} Z^T \underline{E}$$

Considérons l'estimateur  $\hat{\underline{\theta}}$  des paramètres, donné par:

$$\hat{\underline{\theta}} = (Z^T X)^{-1} Z^T \underline{Y} \tag{IV-1}$$

Parmi toutes les matrices Z possibles, celles qui assurent la convergence de l'estimateur  $\hat{\underline{\theta}}$  vers le vecteur  $\underline{\theta}$  des paramètres au sens de la convergence en probabilité<sup>(1)</sup> sont appelées matrices instrumentales.

=====

(1) Convergence en probabilité: La suite des variables aléatoires  $x(n)$  converge en probabilité vers la variable aléatoire  $x$  lorsque  $n$  tend vers l'infini si:

$$\forall \eta_1, \eta_2 > 0 ; \exists N(\eta_1, \eta_2) : n > N \implies P [ |x(n) - x| \geq \eta_1 ] < \eta_2$$

on écrit:  $x(n) \xrightarrow{P} x$

=====

Pour être une matrice instrumentale, Z doit vérifier les deux propriétés suivantes:

$$\frac{1}{M} Z^T \underline{E} \xrightarrow{P} 0 \tag{IV-2}$$

$M \rightarrow \infty$

$$\frac{1}{M} Z^T X \xrightarrow{P} A \tag{IV-3}$$

$M \rightarrow \infty$

A étant une matrice régulière.

Dans ces conditions, l'estimateur  $\hat{\underline{\theta}} = (\underline{Z}^T \underline{X})^{-1} \cdot \underline{Z}^T \underline{Y}$  converge en probabilité vers les vrais paramètres  $\underline{\theta}$ .

En effet :

$$\underline{\theta} = (\underline{Z}^T \underline{X})^{-1} \underline{Z}^T \underline{Y} - (\underline{Z}^T \underline{X})^{-1} \underline{Z}^T \underline{E}$$

donc :

$$\hat{\underline{\theta}} = \underline{\theta} + (\underline{Z}^T \underline{X})^{-1} \underline{Z}^T \underline{E}$$

$$\hat{\underline{\theta}} - \underline{\theta} = (\underline{Z}^T \underline{X})^{-1} \underline{Z}^T \underline{E}$$

$$\hat{\underline{\theta}} - \underline{\theta} \xrightarrow{p} \underline{A}^{-1} \cdot \underline{0} = \underline{0}$$

$$\hat{\underline{\theta}} \xrightarrow{p} \underline{\theta}$$

### UNE CLASSE DE MATRICES INSTRUMENTALES

Soit  $Z$  une matrice de dimensions  $M \times m$  dont les éléments sont bornés :

$$|z_{i,j}| \leq z_{\max}$$

Si les éléments  $z_{i,j}$  sont indépendants des éléments de  $\underline{E}$ , alors :

$$\frac{1}{M} \underline{Z}^T \underline{E} \xrightarrow{p} \underline{0} \quad (\text{IV-4})$$

### IV-2 QUELQUES EXEMPLES DE MATRICES INSTRUMENTALES

#### IV-2-1 Matrice instrumentale X

Considérons l'estimateur des moindres carrés :

$$\hat{\underline{\theta}} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{Y}$$

On remarque que  $X$  joue le rôle de la matrice instrumentale  $Z$ ; cette matrice  $X$  n'assure la convergence en probabilité de l'estimateur vers les paramètres que si les éléments de  $X$  et de  $\underline{E}$  sont indépendants, ce qui n'est généralement pas le cas.

#### IV-2-2 Matrice instrumentale à estimateur de MARKOV

Rappelons l'estimateur de MARKOV :

$$\hat{\underline{\theta}} = (\underline{X}^T \underline{R}^{-1} \underline{X})^{-1} \underline{X}^T \underline{R}^{-1} \underline{Y}$$

où R est la matrice de covariance des résidus. Ici, la matrice instrumentale choisie est :

$$Z = R^{-1} X$$

Pour les mêmes raisons de convergence que précédemment, la matrice X et le vecteur  $\underline{E}$  ne doivent pas être corrélés.

#### IV-2-3 Matrices instrumentales optimales

On peut poser :

$$X = X_d + X_a$$

- $X_d$  : partie déterministe de X, formée à partir de  $v(t), u(t)$

- $X_a$  : partie aléatoire de X, où se trouvent tous les éléments dûs au bruit. On démontre que  $X_d$  est une matrice instrumentale et que la matrice :

$$Z = R^{-1} X_d$$

dans laquelle R est la covariance des résidus, est optimale en ce sens qu'elle minimise la variance de l'estimateur. L'estimateur que donne cette méthode est celui de MARKOV, dans lequel on a remplacé X par  $X_d$ , pour assurer la convergence de cet estimateur. Cette méthode ne peut pourtant pas recevoir d'application pratique du fait qu'elle nécessite la connaissance de  $X_d$ , laquelle est inaccessible.

Certains auteurs proposent de prendre la matrice suivante :

$$Z = \begin{bmatrix} -y_m^{(N-1)} & \dots & -y_m^{(N-n)} & u(N) & \dots & u(N-p) \\ -y_m^{(N-2)} & \dots & -y_m^{(N-n-1)} & u(N-1) & \dots & u(N-p-1) \\ \vdots & & & & & \\ -y_m^{(n)} & \dots & -y_m^{(1)} & u(n+1) & \dots & u(n+1-p) \end{bmatrix} \quad (IV-5)$$

où  $y_m(t)$  est la sortie du modèle à l'instant t :

$$y_m(t) = - \sum_{i=1}^n a_i y_m(t-i) + \sum_{i=0}^p b_i u(t-i)$$

Les  $\hat{a}_i$  et  $\hat{b}_i$  sont les paramètres estimés par un calcul antérieur, la méthode devant être employée de façon itérative. Au fur et à mesure que ces paramètres tendent vers les vrais paramètres  $a_i$  et  $b_i$ ,  $y_m(t)$  tend vers  $v(t)$  et  $Z$  vers  $X_d$ .

On suggère alors :

- d'initialiser le calcul par une méthode des moindres carrés,
- d'utiliser les paramètres obtenus pour calculer  $y_m$  aux divers instants d'échantillonnage,
- et enfin de calculer un nouvel estimateur par la méthode de la matrice instrumentale.

Cette démarche est surtout utilisée en temps réel.

#### IV-2-4 Matrice instrumentale de BANON

BANON a proposé la matrice instrumentale suivante:

$$X = \begin{bmatrix} -y^{(N-1)} & \dots & -y^{(N-n)} & u^{(N)} & \dots & u^{(N-p)} \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \end{bmatrix}$$

alors:

$$Z = \begin{bmatrix} -y^{(N-1-K)} & \dots & -y^{(N-n-K)} & u^{(N)} & \dots & u^{(N-p)} \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \end{bmatrix} \quad (\text{IV-6})$$

avec  $K = n$

Cette matrice est dite à observations retardées. Une partie de ses éléments -les mesures des sorties bruitées- comportent un retard par rapport aux éléments correspondants de  $X$ . L'introduction de ce retard a pour but la décorrélation entre les deux matrices  $Z$  et  $\underline{E}$ .

IV-2-5 Matrice instrumentale de BANON modifiée

Mme S. GENTIL, dans son étude (Etude comparative de diverses méthodes statistiques d'identification; thèse de 3e cycle, Grenoble, 1972), qui a servi de base à ce présent projet, a proposé la modification suivante : on choisit  $K \geq n$  ; la raison de ce choix est expliquée ainsi :

Le résidu  $e(t)$  étant corrélé, on peut prendre pour modèle générateur de bruit :

$$e(t) = G(z^{-1}) \xi(t) = \xi(t) + g_1 \xi(t-1) + \dots + g_q \xi(t-q) \quad (IV-7)$$

où  $\xi(t)$  est un bruit blanc.

Pour que la matrice  $Z$  de l'équation (IV-6) vérifie la propriété (IV-4), il faut que  $y(N-1-K), \dots, y(N-n-K)$  soient indépendants de  $e(N)$ , cette remarque étant valable pour toutes les lignes de  $Z$ . Ceci est vérifié si :

$$K \geq q$$

or: 
$$e(t) = A(z^{-1}) \cdot b(t) \quad ( b(t): \text{bruit additif de sortie} )$$

Le modèle choisi pour les résidus étant celui donné par (IV-7), on déduit que l'on a :

$$q \geq n$$

avec égalité si le bruit  $b(t)$  est blanc.

On a donc :

$$K \geq q \text{ et } q \geq n \text{ -----} \rightarrow K \geq n$$

REMARQUE : On ignore la valeur à donner à  $K$  ; pour l'obtenir, il est nécessaire d'en essayer plusieurs; pour cela, on se basera sur un critère pour déterminer la valeur optimale de ce paramètre. Nous reviendrons sur ce choix dans

le chapitre VII , consacré aux résultats pratiques. D'autre part, si dans la méthode des moindres carrés, n valeurs entrée-sorties suffisaient pour initialiser la récurrence, il en faut ici n+K; donc N données permettent d'écrire N-n-K équations. Pour perdre le moins d'information possible, il est donc souhaitable de donner une faible valeur à K.

Pour ce qui est de la propriété (IV-3),

$$\frac{1}{M} \cdot Z^T X \xrightarrow{P} A, A \text{ régulière,}$$

BANON montre que l'existence de M est conditionnée à la fois par l'entrée et par les paramètres à déterminer. Il n'est pas possible alors d'énoncer une règle générale.

#### IV-3 ESTIMATION DE LA VARIANCE DES RESIDUS =====

L'estimateur est :

$$\hat{\underline{E}} = \underline{Y} - X \hat{\underline{\theta}}$$

sa convergence en probabilité est assurée car :

$$\hat{\underline{E}} = \underline{Y} - X \left[ \underline{\theta} + (Z^T X)^{-1} Z^T \underline{E} \right] = \underline{E} - X (Z^T X)^{-1} Z^T \underline{E}$$

avec :

$$\frac{1}{M} Z^T \underline{E} \xrightarrow{P} \underline{0}$$

et :

$$\frac{1}{M} Z^T X \xrightarrow{P} A$$

donc

$$\hat{\underline{E}} \xrightarrow{P} \underline{E}$$

Par ailleurs:

$$\begin{aligned} \hat{\sigma}_e^2 &= \frac{1}{M} \hat{\underline{E}}^T \hat{\underline{E}} = \frac{1}{M} ( X [\underline{\theta} - \hat{\underline{\theta}}] + \underline{E} )^T ( X [\underline{\theta} - \hat{\underline{\theta}}] + \underline{E} ) \\ &= \frac{1}{M} ( [\underline{\theta} - \hat{\underline{\theta}}]^T X^T X [\underline{\theta} - \hat{\underline{\theta}}] + \underline{E}^T \underline{E} + \underline{E}^T X [\underline{\theta} - \hat{\underline{\theta}}] + [\underline{\theta} - \hat{\underline{\theta}}]^T X^T \underline{E} ) \end{aligned}$$

$$\text{mais : } (\underline{e} - \hat{\underline{e}}) \xrightarrow{P} \underline{0} \quad \text{et} \quad \frac{1}{M} \underline{E}^T \underline{E} \xrightarrow{P} \underline{\sigma}_e^2$$

$$\text{donc : } \hat{\underline{\sigma}}_e^2 \xrightarrow{P} \underline{\sigma}_e^2$$

## B OPTIQUE TEMPS REEL

### IV-4 ALGORITHME D'ESTIMATION DES PARAMETRES

=====

Comme pour les moindres carrés, l'algorithme d'estimation des paramètres découle directement des équations établies en temps différé. En outre, l'estimateur des moindres carrés :

$$\hat{\underline{\theta}} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T \underline{Y}$$

et celui de la matrice instrumentale :

$$\hat{\underline{\theta}} = (\underline{Z}^T \underline{X})^{-1} \underline{Z}^T \underline{Y}$$

ont la même expression formelle,  $\underline{Z}$  différant de  $\underline{X}$  par le retard  $K$ ; l'algorithme pour la matrice instrumentale en temps réel s'obtient par des calculs tout à fait identiques. On aboutit donc à :

$$\begin{aligned} \hat{\underline{\theta}}_{N+1} &= \hat{\underline{\theta}}_N + \underline{K}_{N+1} (y^{(N+1)} - x^{(N+1)} \hat{\underline{\theta}}_N) \\ \underline{K}_{N+1} &= P_N \underline{z}^T(N+1) \left[ 1 + x^{(N+1)} P_N \underline{z}^T(N+1) \right]^{-1} \\ P_{N+1} &= P_N - \underline{K}_{N+1} x^{(N+1)} P_N \end{aligned}$$

où :  $x^{(N+1)}$  et  $z^{(N+1)}$  sont des vecteurs ligne

$$x^{(N+1)} = \left[ -y^{(N)} \dots -y^{(N-n+1)} \quad u^{(N+1)} \dots u^{(N+1-p)} \right]$$

$$z^{(N+1)} = \left[ -y^{(N-K)} \dots -y^{(N-n+1-K)} \quad u^{(N+1)} \dots u^{(N+1-p)} \right]$$

La décorrélation de  $z(N+1)$  avec l'erreur de prédiction du modèle,  $e(N+1)$ , est assurée par une valeur de  $K$  à choisir judicieusement.

Dans cet algorithme, on a posé  $P_N = (Z_N^T X_N)^{-1}$ ; cette matrice pourra être initialisée comme dans la méthode des moindres carrés.

L'avantage de cette méthode par rapport à la précédente est qu'elle n'exige pas de supposer  $e(t)$  blanc pour assurer la convergence de l'estimateur, d'autant plus que cette hypothèse-pour le modèle choisi- n'est jamais vérifiée.

IV- ESTIMATION DE LA VARIANCE DES RESIDUS EN TEMPS REEL  
 =====

L'estimateur  $\hat{\sigma}_e^2$  de la variance des résidus a été calculé plus haut ; on a :

$$\hat{\sigma}_{eN}^2 = \frac{1}{M - m} \cdot \hat{E}_N^T \cdot \hat{E}_N \quad \text{avec } N \text{ mesures}$$

$M$  : nombre d'équations écrites avec  $N$  mesures

$m = n+p+1$

$\hat{E}_N$  : estimateur des résidus à l'instant  $N$

$$\hat{E}_N = Y_N - X_N \hat{\theta}_N$$

Les données arrivent au fur et à mesure que leur prélèvement s'effectue et ne sont pas enregistrées. L'estimation de la variance  $\hat{\sigma}_e^2$  doit se faire de façon récurrente.

A l'étape suivante  $N+1$  nous aurons donc :

$$\begin{aligned} \hat{\sigma}_{eN+1}^2 &= \frac{1}{M+1 - m} ( Y_{N+1} - X_{N+1} \hat{\theta}_{N+1} )^T ( Y_{N+1} - X_{N+1} \hat{\theta}_{N+1} ) \\ &= \frac{1}{M + 1 - m} ( Y_{N+1}^T Y_{N+1} - 2 Y_{N+1}^T X_{N+1} \hat{\theta}_{N+1} + \hat{\theta}_{N+1}^T X_{N+1}^T X_{N+1} \hat{\theta}_{N+1} ) \end{aligned}$$

$$\hat{\sigma}_{eN+1}^2 = \frac{1}{M+1-m} \left[ \underline{y}_N^T \underline{y}_N + y^{2(N+1)} - 2 \left( \underline{y}_N^T \underline{x}_N + y^{(N+1)} \underline{x}^{(N+1)} \right) \cdot \hat{e}_{N+1} \right. \\ \left. + \hat{e}_{N+1}^T \left( \underline{x}_N^T \underline{x}_N + \underline{x}^T(N+1) \underline{x}^{(N+1)} \right) \hat{e}_{N+1} \right]$$

D'où l'algorithme final :

$$\hat{\sigma}_{eN+1}^2 = \frac{1}{M+1-m} \left( \alpha_{N+1} - 2 \beta_{N+1} \hat{e}_{N+1} + \hat{e}_{N+1}^T \Gamma_{N+1} \hat{e}_{N+1} \right)$$

$$\text{avec } \begin{cases} \alpha_{N+1} = \alpha_N + y^{2(N+1)} \\ \beta_{N+1} = \beta_N + y^{(N+1)} \underline{x}^{(N+1)} \\ \Gamma_{N+1} = \Gamma_N + \underline{x}^T(N+1) \underline{x}^{(N+1)} \end{cases}$$

Son initialisation se fait comme suit :

$$\alpha_0 = 0 \quad ; \quad \beta_0 = 0 \quad ; \quad \Gamma_0 = 0$$

CHAPITRE V

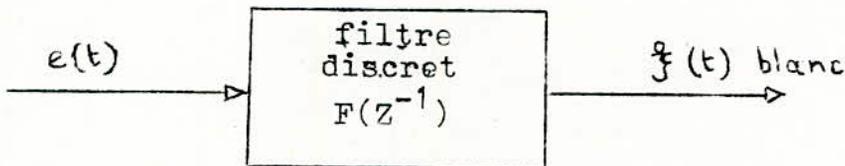
M E T H O D E D E S  
M O I N D R E S C A R R E S G E N E R A L I S E S

Nous avons vu au chapitre III que la méthode des moindres carrés simples présente un gros avantage pour l'identification dynamique: sa mise en oeuvre simple et la forme explicite de l'estimateur qu'elle fournit. Mais celui-ci, du fait que les résidus ne sont jamais blancs, est un estimateur biaisé. Pour pallier à cet inconvénient, tout en conservant à la méthode sa simplicité, plusieurs chercheurs ont proposé un moyen de rendre ces résidus blancs, par transformation des données c'est la technique du filtrage.

V-1 METHODES DE FILTRAGE

V-1-1 Principe

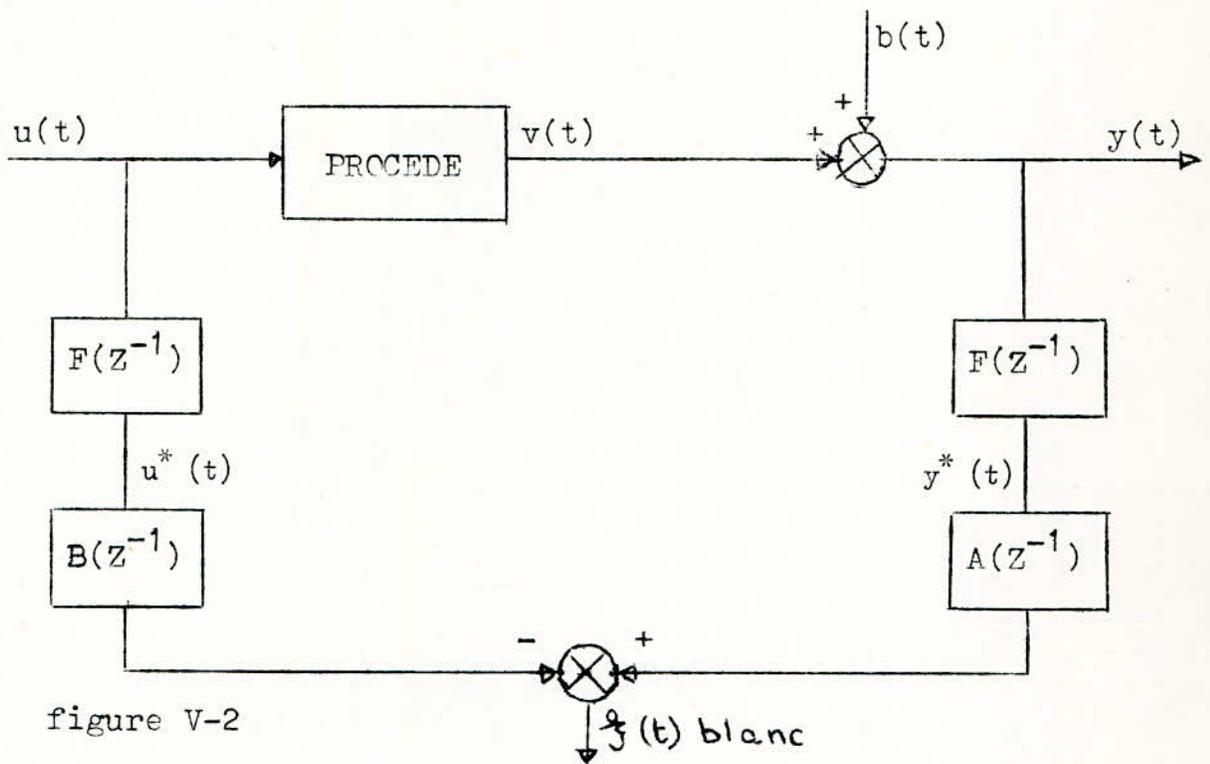
Il s'agit de déterminer un filtre qui "blanchisse" les résidus, l'entrée de ce filtre étant  $e(t)$ , la sortie devant être un bruit blanc  $\xi(t)$



$$F(Z^{-1}) e(t) = \xi(t)$$

Il convient ensuite de filtrer les données par ce même filtre, ce qui revient à filtrer l'équation:  $\underline{Y} = X \underline{\theta} + \underline{E}$

On obtient ainsi un problème des moindres carrés simples ayant un résidu blanc  $\xi(t)$ , donc un estimateur non biaisé. La figure V-2 schématise ce principe.



La transformation des équations est donc la suivante:

$$A(Z^{-1}) y(t) = B(Z^{-1}) u(t) + e(t) \text{ (moindres carrés)}$$

Après filtrage, nous obtenons:

$$A(Z^{-1}) F(Z^{-1}) y(t) = B(Z^{-1}) [F(Z^{-1}) u(t)] + [F(Z^{-1}) e(t)]$$

$$\implies A(Z^{-1}) y^*(t) = B(Z^{-1}) u^*(t) + \xi(t)$$

#### V-1-2 Filtre idéal

Considérons l'estimateur de MARKOV (cf I-2-2):

$$\hat{\underline{\theta}} = (\underline{X}^T \underline{R}^{-1} \underline{X})^{-1} \underline{X}^T \underline{R}^{-1} \underline{Y}$$

où  $\underline{R} = E.M. \{ \underline{E}. \underline{E}^T \}$  est la matrice de covariance des résidus. Par construction,  $\underline{R}$  est une matrice symétrique, définie positive.

Elle peut se décomposer en un produit de deux matrices triangulaires, régulières :

$$R = Q Q^T$$

Le filtrage des données s'opère par  $Q^{-1}$  :

$$\begin{aligned} \underline{Y} &= X \underline{\theta} + \underline{E} \\ \implies Q^{-1} \underline{Y} &= Q^{-1} X \underline{\theta} + Q^{-1} \underline{E} \\ \underline{Y}^* &= X^* \underline{\theta} + \underline{E}^* \end{aligned}$$

L'estimateur de  $\underline{\theta}$  s'écrit donc, après filtrage :

$$\begin{aligned} \hat{\underline{\theta}} &= (X^{*T} X^*)^{-1} X^{*T} \underline{Y}^* \\ \hat{\underline{\theta}} &= (X^T Q^{-1T} Q^{-1} X)^{-1} X^T Q^{-1T} Q^{-1} \underline{Y} \\ \hat{\underline{\theta}} &= (X^T R^{-1} X)^{-1} X^T R^{-1} \underline{Y} \end{aligned}$$

Il est à variance minimale (estimateur de MARKOV). Par ailleurs :

$$\begin{aligned} E.M \{ \underline{E}^* \underline{E}^{*T} \} &= E.M \{ Q^{-1} \underline{E} \underline{E}^T Q^{-1T} \} \\ &= Q^{-1} E.M \{ \underline{E} \underline{E}^T Q^{-1T} \} \\ &= Q^{-1} R Q^{-1T} \\ &= Q^{-1} Q Q^T Q^{-1T} = [I] \end{aligned}$$

Le filtrage des données conduit à des résidus blancs, donc à un estimateur sans biais. Mais ce filtrage par  $Q^{-1}$  est irréalisable, car on ne peut connaître la covariance des résidus, celle-ci étant déterminée avec la covariance du bruit additif de sortie  $b(t)$  et les paramètres  $a_i, b_i$  du procédé qui sont eux-mêmes inconnus ( $e(t) = A(Z^{-1}) b(t)$ ).

V-1-3 Méthode de STEIGLITZ et Mc BRIDE

A partir du critère prédictif des moindres carrés  $C = \sum_{n+1}^N e^2(t)$ , STEIGLITZ et Mc BRIDE proposent un passage au critère quadratique :

$$C_M = \sum_{t=n+1}^N \left[ y(t) - \frac{B(Z^{-1})}{A(Z^{-1})} u(t) \right]^2$$

où  $\frac{B(Z^{-1})}{A(Z^{-1})} u(t) = y_m(t)$  représente la sortie du modèle et  $y(t)$

la sortie du procédé. Si ce dernier est bien identifié,  $A(Z^{-1})$  et  $B(Z^{-1})$  sont identiques pour le modèle et le procédé. On aurait alors :

$$y_m(t) = v(t)$$

Le modèle ainsi obtenu fournit une sortie non bruitée qui est la sortie déterministe  $v(t)$  du procédé, non affectée par le bruit additif  $b(t)$ . Dans ce cas :

$$C_M = \sum_{t=n+1}^N [y(t) - v(t)]^2 = \sum_{n+1}^N b^2(t)$$

La démarche proposée pour passer de  $C$  à  $C_m$  est la suivante :

- une première estimation par les moindres carrés donne les polynômes  $A_1(Z^{-1})$  et  $B_1(Z^{-1})$
- on filtre les données par  $1/A_1(Z^{-1})$  :

$$y^*(t) = \frac{y(t)}{A_1(Z^{-1})}$$

$$u^*(t) = \frac{u(t)}{A_1(Z^{-1})}$$

Le critère prédictif  $C$  devient :

$$\begin{aligned} C' &= \sum_{t=n+1}^N \left[ A_1(Z^{-1}) y^*(t) - B_1(Z^{-1}) u^*(t) \right]^2 \\ &= \sum_{t=n+1}^N \left[ \frac{A_1(Z^{-1})}{A_1(Z^{-1})} y(t) - \frac{B_1(Z^{-1})}{A_1(Z^{-1})} u(t) \right]^2 \end{aligned}$$

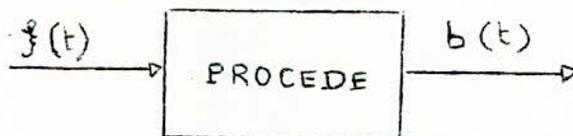
On aura donc  $C' = C_m$  si  $A(Z^{-1}) = A_1(Z^{-1})$ , c'est-à-dire si la première estimation est convenable; sinon, on itère ce processus de filtrage de façon à avoir:

$$C \dashrightarrow C_m \quad \text{ou} \quad \sum_{t=n+1}^N e^2(t) \text{ tendant vers } \sum_{t=n+1}^N b^2(t)$$

Pour que cette approche ait un intérêt dans la solution des moindres carrés, il faut-pour obtenir un estimateur sans biais-que  $b(t)$  soit blanc. Cette méthode serait adéquate pour blanchir les résidus. Mais dans les processus industriels, il est rare que  $b(t)$  soit blanc et l'efficacité de cette méthode s'en trouve diminuée.

V-1-4 Méthode de CLARKE : moindres carrés généralisés

Le point de départ de cette méthode est la représentation du bruit de sortie  $b(t)$  sous forme d'un signal issu d'un bruit blanc. Ceci peut être obtenu en excitant un procédé par un bruit blanc. On obtient alors en sortie  $b(t)$  : le procédé joue alors le rôle de générateur de bruit corrélé.



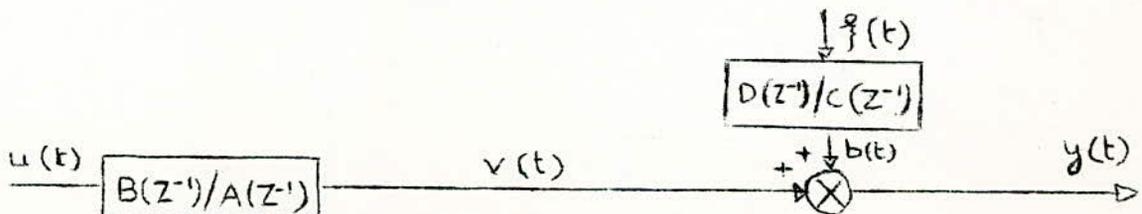
fig( V-3)

D'une façon générale, le générateur de bruit corrélé a une fonction de transfert de la forme:

$$\frac{b(t)}{f(t)} = \frac{D(Z^{-1})}{C(Z^{-1})} = \frac{\sum_0^p d_i Z^{-i}}{1 + \sum_1^n c_{i-1} Z^{-i}}$$

avec :  $p \leq n$  et les racines de  $C(Z^{-1})$  intérieures au cercle unité (condition de stabilité)

Le système figure(II-1) prend la forme suivante:



Les résidus s'écrivent alors dans la forme suivante:

$$e(t) = A(z^{-1}) b(t)$$

$$e(t) = A(z^{-1}) \frac{D(z^{-1})}{C(z^{-1})} \mathfrak{f}(t)$$

L'équation décrivant le fonctionnement du système

$$A(z^{-1}) y(t) = B(z^{-1}) u(t) + e(t) \quad \text{devient:}$$

$$A(z^{-1}) y(t) = B(z^{-1}) u(t) + A(z^{-1}) \frac{D(z^{-1})}{C(z^{-1})} \mathfrak{f}(t)$$

On voit dès lors que le filtre de blanchiment théorique doit être:

$$F(z^{-1}) = \left[ A(z^{-1}) \frac{D(z^{-1})}{C(z^{-1})} \right]^{-1}$$

REMARQUE : Dans la méthode de STEIGLITZ et Mc BRIDE on prenait:  
=====

$$F(z^{-1}) = \left[ A(z^{-1}) \right]^{-1}$$

ce qui impliquait  $\frac{C(z^{-1})}{D(z^{-1})} = 1$  donc que  $b(t) = \mathfrak{f}(t)$ :

le bruit sur le système devrait être blanc pour avoir un filtrage idéal.

L'approche de CLARKE est plus générale puisqu'elle ne suppose pas  $b(t)$  blanc. CLARKE propose, pour déterminer expérimentalement ce filtre  $F(z^{-1})$ , de l'approcher par un développement en série d'ordre  $q$ :

$$F(z^{-1}) = 1 + f_1 z^{-1} + \dots + f_q z^{-q}$$

Le modèle des résidus:

$$F(z^{-1}) e(t) = \mathfrak{f}(t)$$

soit:  $e(t) + f_1 e(t-1) + \dots + f_q e(t-q) = \mathfrak{f}(t) \quad (V-1)$

On l'appelle modèle autorégressif et le filtre correspondant  $F(z^{-1})$  est dit filtre autorégressif.

Aux différents instants d'échantillonnage, on a :

$$\begin{aligned}
 e(N) + f_1 e(N-1) + \dots + f_q e(N-q) &= \xi(N) \\
 e(N-1) + f_1 e(N-2) + \dots + f_q e(N-q-1) &= \xi(N-1) \\
 \vdots & \\
 \vdots & \\
 e(q+1) + f_1 e(q) + \dots + f_q e(1) &= \xi(q+1)
 \end{aligned}$$

$$\begin{aligned}
 \begin{matrix} \text{-----} \Rightarrow \\ \left[ \begin{array}{c} e(N) \\ e(N-1) \\ \vdots \\ e(q+1) \end{array} \right] \end{matrix} &= \begin{matrix} \left[ \begin{array}{ccc} -e(N-1) & \dots & -e(N-q) \\ -e(N-2) & \dots & -e(N-q-1) \\ \vdots & & \vdots \\ -e(q) & \dots & -e(1) \end{array} \right] \end{matrix} \begin{matrix} \left[ \begin{array}{c} f_1 \\ f_2 \\ \vdots \\ f_q \end{array} \right] \end{matrix} + \begin{matrix} \left[ \begin{array}{c} \xi(N) \\ \xi(N-1) \\ \vdots \\ \xi(q+1) \end{array} \right] \end{matrix}
 \end{aligned}$$

$$\text{-----} \Rightarrow \underline{E} = X_E \underline{f} + \underline{\xi}$$

La solution de ce problème est justement celle des moindres carrés :

$$\hat{\underline{f}} = (X_E^T X_E)^{-1} X_E^T \underline{E}$$

Comme  $\xi(t)$  est blanc, l'estimateur  $\hat{\underline{f}}$  du filtre est sans biais. Ainsi, lorsque les résidus sont connus (par estimation), il est facile de déterminer ce filtre.

Résumé de la méthode  
 =====

i) Un premier calcul fournit l'estimateur des moindres carrés

$$\hat{\underline{\theta}}_1 = (X^T X)^{-1} X^T \underline{Y}$$

ii) On estime les résidus :

$$\hat{\underline{E}}_1 = \underline{Y} - X \hat{\underline{\theta}}_1$$

et on construit la matrice  $X_{E1}$  qui nous permet de calculer le filtre :

$$\hat{\underline{f}}_1 = (X_{E1}^T X_{E1})^{-1} X_{E1}^T \hat{\underline{E}}_1$$

iii) On filtre les données

$$y^*(t) = y(t) + \hat{f}_1 y(t-1) + \dots + \hat{f}_q y(t-q)$$

$$u^*(t) = u(t) + \hat{f}_1 u(t-1) + \dots + \hat{f}_q u(t-q)$$

et on considère les ensembles  $y^*(t)$  et  $u^*(t)$  comme les données d'un nouveau problème. On calcule donc :

$$\hat{e}_2 = (X^{*T} X^*)^{-1} X^{*T} \underline{Y}^*$$

et on itère .

REMARQUE : On suggère en général de prendre  $q = 1$  pour deux raisons :

1) après  $s$  itérations, le filtrage global est  $F(Z^{-1}) = \prod_{i=1}^s F_i(Z^{-1})$ ; l'ordre de ce filtre global est  $q.s$ . On évite une contrainte sur la parité de l'ordre de ce filtre en prenant  $q = 1$ .

2) il faut à chaque itération initialiser l'équation (V-1) avec  $q$  valeurs entrée-sortie pour ce filtrage. On perd ainsi  $q$  valeurs à chaque itération. En prenant  $q = 1$ , on réduit cette perte d'information.

V-2 AUTRE METHODE D'OBTENTION DE LA SOLUTION DES M.C.G. :

MODELE PARAMETRIQUE ELARGI

Dans le cas d'un bruit additif de sortie blanc, l'équation aux différences décrivant le système étant :

$$y(t) + \sum_{i=1}^n a_i y(t-i) = \sum_{i=0}^p b_i u(t-i) + b(t) + \sum_{i=1}^s a_i b(t-i),$$

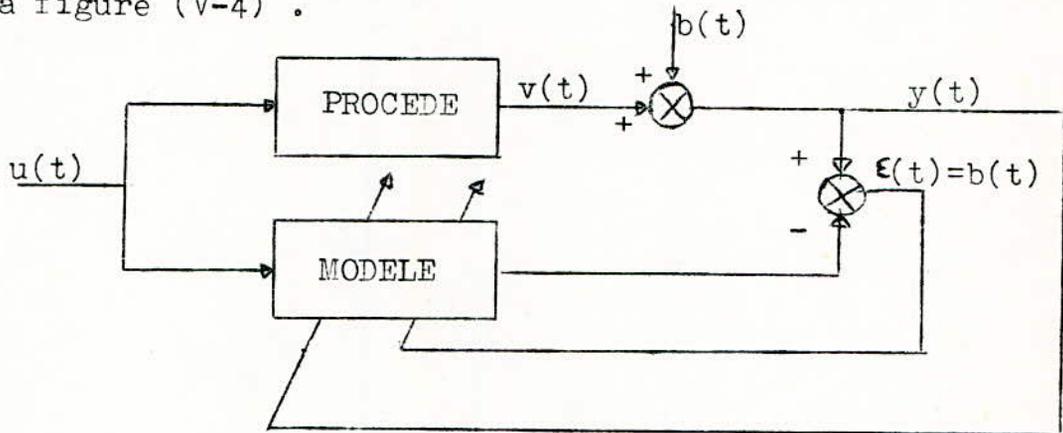
on convient de la considérer sous la forme :

$$y(t) - \left[ - \sum_{i=1}^n a_i y(t-i) + \sum_{i=0}^p b_i u(t-i) + \sum_{i=1}^s a_i b(t-i) \right] = b(t)$$

Le modèle choisi dans ce cas, prédit une sortie  $\hat{y}(t/t-1)$  :

$$\hat{y}(t/t-1) = - \sum_{i=1}^n a_i y(t-i) + \sum_{i=0}^p b_i u(t-i) + \sum_{i=1}^n a_i b(t-i)$$

Le résidu d'un tel modèle (son erreur de prédiction) est  $b(t)$ , supposé blanc. Le schéma d'un tel modèle est représenté sur la figure (V-4) .



Sous forme vectorielle nous avons les mêmes équations que dans le cas des moindres carrés simples ,c'est-à-dire:

$$\underline{Y} = X \underline{\theta} + \underline{B}$$

Les dimensions de X et  $\underline{\theta}$  ne sont plus  $m=n+p+1$  ,mais  $m' = 2n+p+1$  .

$$\underline{\theta}^T = [a_1 \dots a_n \ b_0 \dots b_p \ a_1 \dots a_n]$$

$$X = \begin{bmatrix} -y(N-1) & \dots & -y(N-n+1) & u(N) & \dots & u(N+1-p) & u(N-p) & b(N-1) & \dots & b(N-n+1) \\ -y(N-2) & \dots & & u(N-1) & & & & b(N-2) & & \\ \vdots & & & \vdots & & & & \vdots & & \\ \vdots & & & \vdots & & & & \vdots & & \end{bmatrix}$$

L'estimateur obtenu est ainsi sans biais, avec l'hypothèse de  $b(t)$  blanc :

$$\hat{\underline{\theta}} = (X^T X)^{-1} X^T \underline{Y}$$

Le critère prédictif minimisé est ici:

$$C = \sum_{t=n+1}^N b^2(t)$$

:c'est le critère auquel conduit la méthode de STEIGLITZ et Mc BRIDE. La différence est qu'ici on ne filtre pas les données, on élargit le vecteur des paramètres en lui adjoignant une partition de  $a_i$  ( $i=1, n$ ). On identifie ainsi deux fois les mêmes paramètres, ce qui est un inconvénient. Mais on a réussi à éviter la forte corrélation introduite par l'intermédiaire du polynôme  $A(Z^{-1})$  dans le modèle des moindres carrés:

$$e(t) = A(Z^{-1}) b(t)$$

On peut remarquer que le modèle de la figure(V-4) est un modèle ajustable, excité à la fois par l'entrée et la sortie du procédé et par l'erreur de prédiction

On notera également qu'il est possible de l'utiliser pour estimer la composante, à tout instant, du bruit de sortie  $b(t)$ . Ceci peut être intéressant pour évaluer le niveau de bruit sur certains processus.

# B OPTIQUE TEMPS REEL

## V-3 FILTRE DE KALMAN APPLIQUE A L'IDENTIFICATION

Le filtre de KALMAN décrit dans [B1] et [B5] est une approche basée sur l'emploi des variables d'état. Son application à l'identification fournit un algorithme de prédiction et d'estimation des paramètres en temps réel.

### V-3-1 Cas général de bruit corrélé

On prend pour modèle des résidus un filtre autoregressif:

$$e(K) + \sum_{i=1}^q f_i e(K-i) = \xi(K), \quad \xi(K) \text{ blanc}$$

(c'est le modèle utilisé par CLARKE)

Ce modèle mis sous forme de modèle d'état devient :

$$\underline{E}_K = \begin{bmatrix} e(K-q+1) \\ \vdots \\ e(K) \end{bmatrix}$$

$$\underline{E}_{K+1} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 1 \\ -f_q & -f_{q-1} & \dots & -f_1 & 0 \end{bmatrix} \underline{E}_K + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \xi(K+1) \end{bmatrix}$$

$\underline{\theta}$  étant le vecteur des paramètres à obtenir, considéré (tel  $\underline{E}_K$ ) comme un vecteur d'état; on pose:

$$\underline{Z}_K = \begin{bmatrix} \underline{\theta}_K \\ \underline{E}_K \end{bmatrix}$$

Le système généralisé qui décrit le processus est alors:

$$\left\{ \begin{array}{l} \underline{Z}_{K+1} = \begin{bmatrix} \underline{\theta}_{K+1} \\ \underline{E}_{K+1} \end{bmatrix} = \begin{bmatrix} 1 & & & & 0 \\ & 0 & 1 & 0 & \dots & 0 \\ & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & & -f_q & & & -f_1 \end{bmatrix} \underline{Z}_K + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \xi(K+1) \end{bmatrix} \\ \underline{Y}_K = \begin{bmatrix} x(K) & 0 & 0 & \dots & 1 \end{bmatrix} \underline{Z}_K = y(K) \quad (\text{scalaire}) \end{array} \right.$$

$\underline{Y}_K$  est le vecteur des sorties

$$\underline{x}(K) = [-y(K-1) \dots -y(K-n) \ u(K) \dots u(K-p)]$$

Sous forme condensée :

$$\begin{cases} \underline{Z}_{K+1} = A \underline{Z}_K + \underline{W}_K \\ \underline{Y}_K = C_K \underline{Z}_K \end{cases} \quad (A \text{ est constante})$$

$\underline{W}_K$  est le vecteur de bruit sur le processus. Sa matrice de covariance est:

$$Q = E.M \{ \underline{W} \underline{W}^T \} = \begin{bmatrix} 0 & & 0 \\ \vdots & & \vdots \\ 0 & & \sigma_w^2 \end{bmatrix}$$

KALMAN a proposé un algorithme pour le cas général du filtre qu'il a étudié. Dans le cas qui nous intéresse, cet algorithme prend une forme plus simple, qui est la suivante:

$$\begin{aligned} \text{a) Prédiction à l'instant } K & \begin{cases} \underline{\Theta}_{K/K-1} = A \underline{\Theta}_{K-1/K-1} \\ P_{K/K-1} = A P_{K-1/K-1} + Q \end{cases} \\ \text{b) Estimation à l'instant } K & \begin{cases} \underline{\Theta}_{K/K} = \underline{\Theta}_{K/K-1} + K_K [y(K) - C_K \cdot \underline{\Theta}_{K/K-1}] \\ P_{K/K} = P_{K/K-1} - K_K C_K P_{K/K-1} \\ K_K = P_{K/K-1} C_K^T [C_K P_{K/K-1} C_K^T]^{-1} \end{cases} \quad (V-2) \end{aligned}$$

$P_{K/K-1}$  et  $P_{K/K}$  sont les matrices de covariances de l'état  $\underline{Z}_K$ , définies par KALMAN :

$$P_{K/K-1} = E.M \left\{ (\underline{Z}_K - \underline{Z}_{K/K-1}) (\underline{Z}_K - \underline{Z}_{K/K-1})^T \right\}$$

$$P_{K/K} = E.M \left\{ (\underline{Z}_K - \underline{Z}_{K/K}) (\underline{Z}_K - \underline{Z}_{K/K})^T \right\}$$

La solution (V-2) est complexe: il faut à chaque itération effectuer plus de calculs, ce qui est un grand handicap dans l'utilisation en temps réel. L'hypothèse de résidus blancs simplifie quelque peu cet algorithme, mais en contre-partie,

on perd en précision .

V-3-2 Cas de résidu e(t) blanc

Si on suppose e(t) blanc, on a les simplifications suivantes; par rapport au cas précédent:

$$\underline{E}_{K+1} = \underline{E}_K = e(K) = \xi(K) \quad , \text{ blanc.}$$

On pose alors:

$$\underline{z}_K = \underline{\theta}_K \implies \underline{w}_K = \underline{0}$$

Le système d'état décrivant le fonctionnement du processus devient:

$$\begin{cases} \underline{z}_{K+1} = \underline{\theta}_{K+1} = [1] \underline{\theta}_K + \underline{0} \\ \underline{y}_K = y(K) = \underline{x}(K) \underline{\theta}_K + e(K) \end{cases} \quad \text{ou} \quad \begin{cases} \underline{\theta}_{K+1} = \underline{\theta}_K \\ y(K) = \underline{x}(K) \underline{\theta}_K + e(K) \end{cases}$$

Ici  $\underline{\theta}$  est un vecteur d'état invariant dans le temps. La valeur prédite est identique à la valeur estimée à l'état précédent:

$$\underline{\theta}_{K/K-1} = \underline{\theta}_{K-1/K-1} = \underline{\theta}_{K-1}$$

$$P_{K/K-1} = P_{K-1/K-1} = P_{K-1}$$

On obtient l'algorithme suivant:

$$\begin{cases} \underline{\theta}_K = \underline{\theta}_{K-1} + K_K [y(K) - \underline{x}(K) \underline{\theta}_{K-1}] \\ K_K = P_{K-1} \underline{x}^T(K) [\underline{x}(K) P_{K-1} \underline{x}^T(K) + R_K]^{-1} \\ P_K = P_{K-1} - K_K \underline{x}(K) P_{K-1} \end{cases}$$

$R_K$  est la matrice de covariance du bruit de mesure e(K)

$$R_K = E.M \left\{ e(K) e^T(K) \right\} = \sigma^2$$

Si on compare cet algorithme à celui des moindres carrés, on voit qu'ils sont identiques: cela s'explique par le

fait que le filtre de KALMAN détermine un estimateur à variance minimale et c'est aussi le cas pour les moindres carrés quand le résidu est blanc.

V-4 AUTRES METHODES DES MOINDRES CARRÉS GENERALISES  
=====

EN TEMPS REEL  
=====

V-4-1 Application de la méthode de CLARKE

Le filtre de CLARKE est d'une mise en oeuvre plus facile que celui de KALMAN : en temps réel, on filtre les données au fur et à mesure qu'elles arrivent. Il est alors nécessaire d'emmagasiner un certain nombre de ces données pour estimer les résidus et les paramètres  $f_1$  du filtre générateur de bruit. Cela demande beaucoup de place en mémoire. De plus, le filtre  $F(Z^{-1})$  dans le cas du modèle choisi pour cette technique, dépend du polynôme  $A(Z^{-1})$  : il doit donc évoluer en même temps que les paramètres du système. Néanmoins, il est possible de l'estimer en temps réel puisqu'il est calculé par une méthode des moindres carrés.

V-4-2 Cas du modèle paramétrique élargi

La solution de la méthode utilisant un modèle élargi à  $2n+p+1$  paramètres est très facile à mettre en oeuvre en temps réel. L'algorithme utilisé est celui des moindres carrés, à ceci près : les dimensions de  $X$  et  $\underline{\theta}$  sont à ajuster à  $2n+p+1$ . Elle présente un avantage certain sur les autres méthodes vues plus haut quant à sa simplicité. Les calculs ne sont pas longs et l'encombrement en mémoire faible, car il n'y a pas de filtre à estimer.

L'initialisation de l'algorithme est semblable à celle de l'algorithme des moindres carrés. On perd en plus  $n$  données au départ, nécessaires pour remplir le vecteur  $\underline{B}$ , initialisé à  $\underline{0}$ , car les  $b(t)$  sont estimés au cours des itérations, la connaissance des valeurs de départ étant inaccessible.

Mais, comme pour la méthode de STEIGLITZ et Mc BRIDE, la simplicité a été obtenue au détriment de la précision. Nous avons supposé  $b(t)$  blanc au départ: dans le cas de processus fortement bruités, on n'est pas assuré de la convergence vers les vrais paramètres. Toutefois - et l'expérimentation le confirme - les résultats sont meilleurs que ceux donnés par les moindres carrés simples. Pour de nombreux processus industriels, la précision atteinte par cette méthode s'avère largement suffisante.

CHAPITRE VI

LES SYSTEMES ADAPTATIFS ECHANTILLONNES  
ET LEUR UTILISATION  
POUR L'IDENTIFICATION DES PROCÉDES PHYSIQUES

INTRODUCTION

=====

On présente dans ce chapitre une méthode d'identification en temps réel des procédés physiques, fondée sur le concept d'hyperstabilité [ POPOV ] et largement développée dans les travaux de I.D. LANDAU.

Rappelons que le but de l'identification en temps réel est de suivre les variations des paramètres et des variables d'état du procédé physique en cours de fonctionnement, et que le problème peut être formulé comme un problème de poursuite. La difficulté principale réside dans la détermination d'un mécanisme d'adaptation tel que les paramètres du modèle ajustable suivent d'aussi près que possible les variations des paramètres du procédé physique.

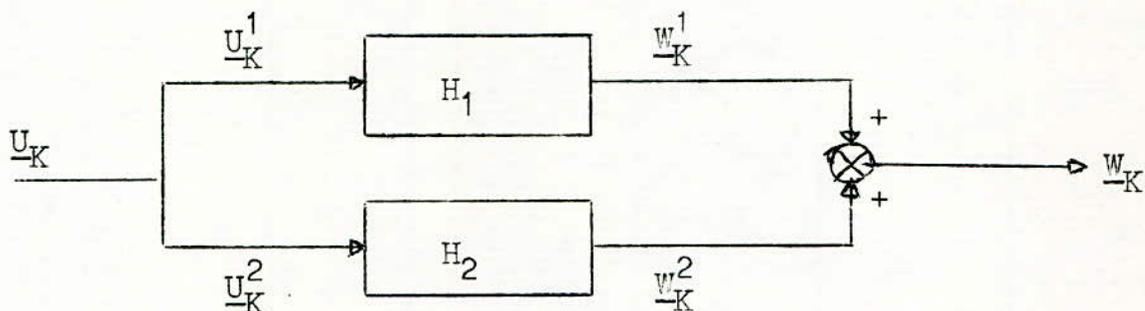
Les méthodes de détermination de tels mécanismes d'adaptation à partir de considérations de stabilité ont très naturellement séduit les automaticiens, plus familiers avec ces techniques qu'avec les méthodes statistiques: en particulier, la deuxième méthode de LIAPOUNOV s'est révélée un outil solide pour un tel travail et a été largement utilisée. Cependant, la détermination d'une fonction de LIAPOUNOV (fonction définie positive telle que sa dérivée soit définie négative sur toute la trajectoire) est quelquefois assez laborieuse.

LANDAU, s'appuyant sur la théorie de l'hyperstabilité développée par POPOV vers 1960, a proposé une méthode d'analyse et de synthèse des systèmes adaptatifs avec modèle de référence, qui permet de décomposer le problème de stabilité du système

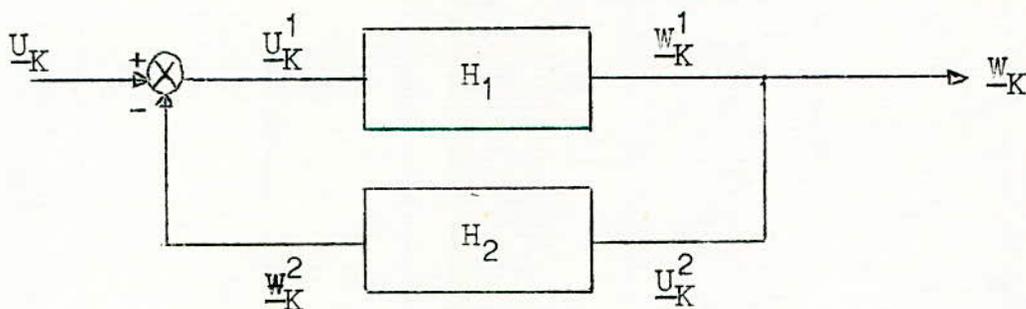
global en plusieurs problèmes d'hyperstabilité de ses composants :

En effet, un système hyperstable peut être caractérisé partiellement par les propriétés suivantes:

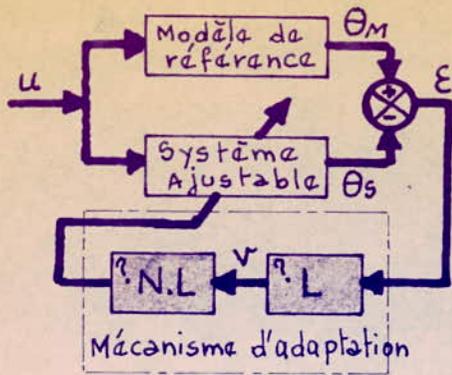
- 1) Stabilité (éventuellement stabilité asymptotique).
- 2) Entrée bornée-sortie bornée.
- 3) La combinaison parallèle de deux systèmes hyperstables est aussi un système hyperstable :



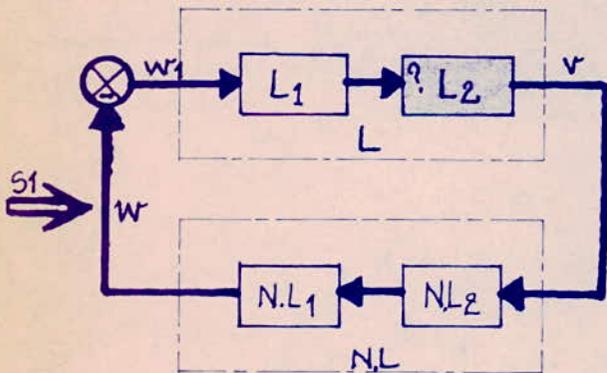
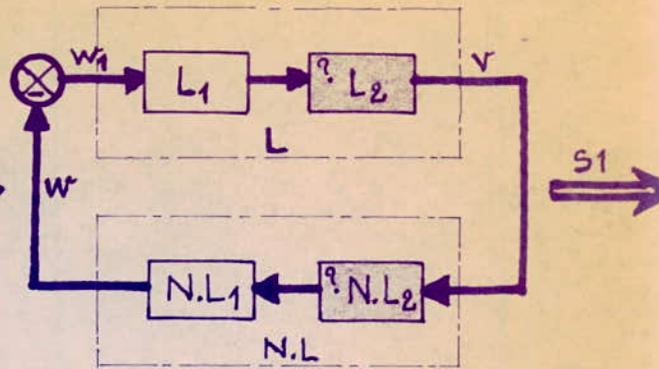
- 4) La combinaison en contre-réaction négative de deux systèmes hyperstables est aussi un système hyperstable:



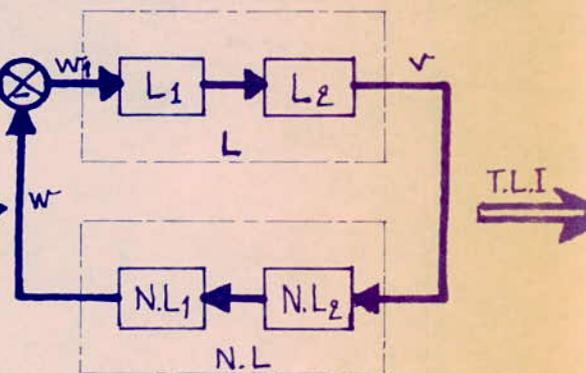
C'est cette dernière propriété qui sera systématiquement utilisée pour la synthèse des systèmes adaptatifs avec modèle de référence.



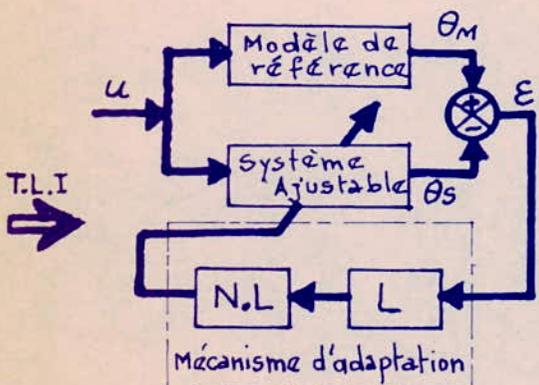
T.L



T.H



T.L.I



Légende

- T.L = transformation linéaire
  - T.L.I = transformation linéaire inverse
  - L = bloc linéaire
  - N.L = bloc non-linéaire
  - T.H = Théorème d'hyperstabilité
  - S1 = solution de l'inégalité intégrale
- $$\int_0^t w^T v dt \geq -\gamma_0^2$$

Diagramme VI-1 : illustration des diverses étapes de la méthode de synthèse permettant de ramener un système adaptatif avec modèle de référence à la forme d'un système équivalent à contre-réaction caractérisé par une partie L et une partie N.L

VI-1 RAPPEL DES CONCEPTS DE BASE

Soit le système H décrit par l'équation en variable d'état discrète :

$$\underline{x}_{K+1} = A_K \underline{x}_K + B_K \underline{u}_K$$

et son équation de sortie :

$$\underline{w}_K = C_K \underline{x}_K$$

où  $\underline{x}_K$ ,  $\underline{u}_K$  et  $\underline{w}_K$  sont des vecteurs

$A_K$ ,  $B_K$  et  $C_K$  des matrices de dimensions appropriées

Considérons aussi la somme des produits entrée-sortie :

$$\mathcal{D}(K_0, K_1) = \sum_{k=K_0}^{K_1} \underline{u}_k^T \underline{w}_k \quad ; \quad K_1 \geq K_0$$

DEFINITION I

H est hyperstable s'il existe deux nombres  $\beta_1 > 0$  et  $\beta_0 \geq 0$  tels que la relation :

$$\mathcal{D}(K_0, K_1) + \beta_0 \|\underline{x}_{K_0}\|^2 \geq \beta_1 \|\underline{x}_{K_1+1}\|^2$$

soit vérifiée.

Interprétation de cette relation: elle représente la généralisation d'une condition de passivité.

$$\underbrace{\sum_{k=K_0}^{K_1} \underline{u}_k^T \underline{w}_k}_{\text{Energie fournie au système sur } [K_0, K_1]} + \underbrace{\frac{1}{2} \underline{x}_{K_0}^T \underline{x}_{K_0}}_{\text{Energie initiale stockée}} \geq \underbrace{\frac{1}{2} \underline{x}_{K_1+1}^T \underline{x}_{K_1+1}}_{\text{Energie stockée a l'instant } K_1+1}$$

(ici  $\beta_0 = \beta_1 = \frac{1}{2}$  et  $\|\underline{x}_K\| = \sqrt{\underline{x}_K^T \underline{x}_K}$  )

DEFINITION II

=====

Si H est hyperstable, on peut trouver un nombre  $\delta > 0$  tel que:

-s'il existe un nombre  $\gamma > 0$  satisfaisant l'inéquation:

$$\eta(K_0, K_1) = \sum_{K=K_0}^{K_1} U_K^T W_K = \gamma^2 + \gamma \sup \|x_K\| \quad \text{pour } K_0 \leq K \leq K_1$$

alors:

$$\|x_K\| \leq \delta (\gamma + \|x_{K_0}\|) \quad \forall K \geq K_0$$

Il en résulte les quatre propriétés suivantes:

Propriété 1 : l'hyperstabilité entraîne la stabilité au sens de LIAPOUNOV (entrée nulle).

Propriété 2 : entrée bornée-sortie bornée.

Propriété 3 : la combinaison parallèle de deux blocs hyperstables est hyperstable.

Propriété 4 : la combinaison en contre réaction négative de deux blocs hyperstables est hyperstable.

DEFINITION III

=====

Un système est asymptotiquement hyperstable si et seulement si :

- 1) il est hyperstable
- 2) à entrée nulle, il est asymptotiquement stable.

DEFINITION IV : HYPERSTABILITE FAIBLE

=====

Considérons un système  $H_W$  décrit par une relation entrée-sortie de la forme:

$$W_K = \mathcal{B}(1, K) U_1 = f(U_1, K) \quad 1 \leq K$$

où  $B(1, K)$  est un opérateur matrice carrée, défini pour tout  $K$ .

Considérons aussi la somme des produits entrée-sortie:

$$\eta(K_0, K_1) = \sum_{K=K_0}^{K_1} U_K^T W_K$$

Le système est dit hyperstable au sens faible s'il existe une constante  $\gamma_0^2 \geq 0$  telle que :

$$\eta(K_0, K_1) = \sum_{K=K_0}^{K_1} U_K^T W_K \geq -\gamma_0^2 \quad ; \quad \forall K_1 \geq K_0$$

REMARQUE

=====

-L'hyperstabilité entraîne toujours l'hyperstabilité au sens faible; l'inverse n'est généralement pas vrai.

-L'hyperstabilité au sens faible est très utilisée pour décrire des blocs non linéaires et/ou non stationnaires.

Propriété 5 : la combinaison parallèle de deux blocs hyperstables au sens faible est hyperstable au sens faible.

Propriété 6 : la combinaison en contre-réaction de deux blocs hyperstables au sens faible est hyperstable au sens faible.

VI-2 APPLICATION A L'IDENTIFICATION

=====

On considère un processus monovarié non bruité, décrit par l'équation :

$$\theta_M(K) = \sum_{l=1}^n a_l \cdot \theta_M^{(K-1)} + \sum_{l=0}^m b_l \cdot \varphi_{K-1}$$

et un modèle ajustable :

$$\theta_S(K) = \sum_{l=1}^n \hat{a}_l(K) \theta_S^{(K-1)} + \sum_{l=0}^m \hat{b}_l(K) \varphi_{K-1}$$

On définit l'erreur entre sortie du processus et sortie du modèle ajustable

$$\varepsilon(K) = \theta_M(K) - \theta_S(K)$$

et la sortie d'un bloc de traitement linéaire de l'erreur:

$$V(K) = \varepsilon(K) + \sum_{\ell=1}^n c_\ell \varepsilon(K-\ell)$$

Les lois d'adaptation des paramètres seront de la forme :

$$\hat{a}_1(K) = \hat{a}_1(K-1) + \varphi_{a_1}(V_K) \quad 1 \leq 1 \leq n$$

$$\hat{b}_1(K) = \hat{b}_1(K-1) + \varphi_{b_1}(V_K) \quad 1 \leq 1 \leq m$$

Le système d'équations précédent peut être représenté sous la forme condensée suivante :

$$\begin{cases} \varepsilon_K = a_M^T e_{K-1} + [p - p(K)]^T y_{K-1} \\ V_K = \varepsilon_K + c^T e_{K-1} \end{cases} \quad (\text{VI-1})$$

où :

$$\hat{p}(K) = \hat{p}(K-1) + \Phi(V_K)$$

$$a_M^T = [a_1 \dots a_n]$$

$$p^T = [a_1 \dots a_n \quad b_1 \dots b_m]$$

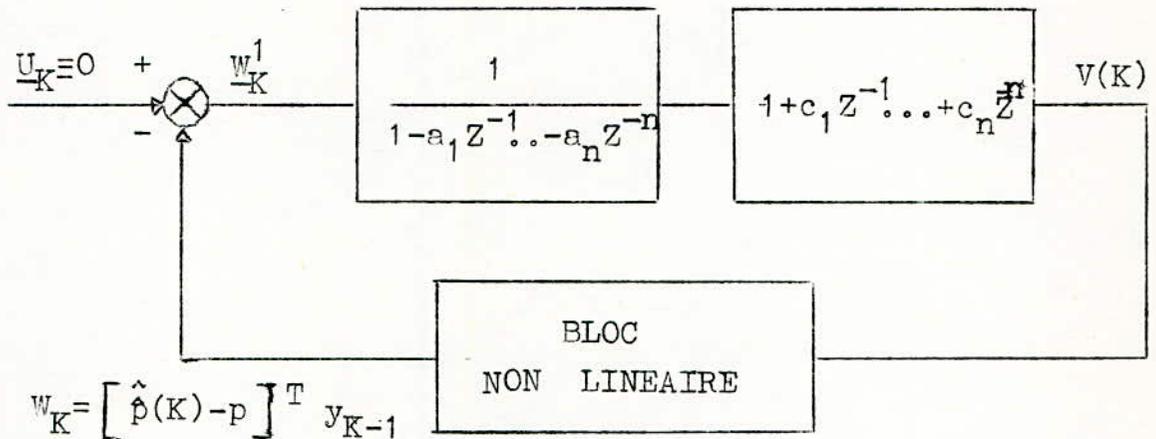
$$\hat{p}^T(K) = [\hat{a}_1(K) \dots \hat{a}_n(K) \quad \hat{b}_1(K) \dots \hat{b}_m(K)]$$

$$y_{K-1}^T = [\theta_S(K-1) \dots \theta_S(K-n) \quad \rho(K-1) \dots \rho(K-m)]$$

$$e_{K-1}^T = [\epsilon_{K-1} \dots \epsilon_{K-n}]$$

$$c^T = [c_1 \dots c_n]$$

Le système d'équations (VI-1) peut donc être représenté sous la forme d'un schéma en contre réaction à entrée nulle:



Les conditions d'hyperstabilité du bloc linéaire de la chaîne directe et du bloc non linéaire de la chaîne de contre-réaction vont nous permettre de déterminer la structure du mécanisme d'adaptation, à savoir:

1) le correcteur de traitement linéaire qui doit être calculé de telle manière que :

$$G(Z) = \frac{1 + \sum_{i=1}^n c_i Z^{-i}}{1 - \sum_{i=1}^n a_i Z^{-i}}$$

soit une fonction de transfert strictement réelle positive. Ainsi, la condition d'hyperstabilité prend cette forme simple pour des blocs linéaires.

2) les lois d'adaptation sont par exemple de la forme:

$$\hat{\Phi}(V_K) = V_K R y_{K-1}$$

où R est une matrice définie positive

En effet, pour ce type d'adaptation, la condition d'hyperstabilité faible est vérifiée:

$$\begin{aligned} \eta(0, K_1) &= \sum_{K=0}^{K_1} V_K^T W_K = \sum_{K=0}^{K_1} V_K^T \left[ \hat{P}_0(0) + \sum_{\ell=0}^K V_\ell R y_{\ell-1} - P \right]^T y_{K+1} \\ &= \sum_{K=0}^{K_1} V_K^T \left[ P_0 + \sum_{\ell=0}^K V_\ell R y_{\ell-1} \right]^T y_{K+1} \end{aligned}$$

avec  $P_0 = \hat{P}_0 - P = \text{cte}$

ceci en vertu de la propriété :

$$\sum_{K=0}^{K_1} x_K \left( \sum_{\ell=0}^K x_\ell + c \right) = \frac{1}{2} \left( \sum_{K=0}^{K_1} x_K + c \right)^2 + \frac{1}{2} \sum_{K=0}^{K_1} x_K^2 - \frac{c^2}{2} \Delta - \frac{c^2}{2}$$

REMARQUE : Cette loi d'adaptation intégrale qui fait intervenir  
 =====  
 une matrice de gain définie positive constante n'est pas tout à fait adéquate pour une identification précise des paramètres du procédé physique; en effet, elle pénalise toutes les mesures de la même manière et on ne peut obtenir des valeurs stables des estimateurs des paramètres en présence de bruit important sur le processus. LANDAU a proposé une nouvelle forme d'algorithme d'identification avec gain variable (sa démonstration se base sur le lemme de POPOV-KALMAN-YAKUBOVICH). On obtient ainsi :

$$\hat{p}(K) = \hat{p}(K-1) + R(K-1) y(K-1) V_K^0$$

avec

$$R(K) = R(K-1) - \frac{R(K-1) y(K-1) y(K-1)^T R(K-1)}{1 + y(K-1)^T R(K-1) y(K-1)}$$

où  $V_K^0$  représente la sortie du correcteur de traitement linéaire de l'erreur calculée avec les estimations des paramètres à l'instant  $K-1$  :

$$V_K^0 = \varepsilon_K^0 + c^T e_{K-1} = \frac{\theta_M^{(K)} - \hat{p}^T(K-1) y_{K-1} + c^T e_{K-1}}{1 + y_{K-1}^T R_{K-1} y_{K-1}}$$



A N N E X E

DETERMINATION DU CORRECTEUR  
POUR UN SYSTEME DU  
DEUXIEME ORDRE

Nous avons vu que la fonction de transfert de la partie linéaire est :

$$G(Z) = \frac{1 + c_1 Z^{-1} + c_2 Z^{-2}}{1 - a_1 Z^{-1} - a_2 Z^{-2}}$$

Pour déterminer les valeurs de  $c_1$  et  $c_2$  qui assurent la condition de stricte réelle positivité, on utilise la transformation bilinéaire  $Z^{-1} = (1-W)/(1+W)$  qui transforme le cercle  $|Z| = 1$  en l'axe imaginaire  $\text{Re}(W) = 0$

Après développement, on obtient :

$$G'(W) = \frac{c'_0 + c'_1 W + c'_2 W^2}{a'_0 + a'_1 W + a'_2 W^2}$$

avec:

$$c'_0 = 1 + c_1 + c_2$$

$$a'_0 = 1 - a_1 - a_2$$

$$c'_1 = 2(1 - c_2)$$

$$a'_1 = 2(1 + a_2)$$

$$c'_2 = 1 + c_2 - c_1$$

$$a'_2 = 1 + a_1 - a_2$$

Pour que, dans l'espace de LAPLACE,  $G'(p)$  soit strictement réelle positive, il faut et il suffit que :

- 1)  $G'(W)$  ne soit pas dégénérée
- 2) les pôles de  $G'(W)$  soient situés dans la moitié gauche du plan  $p$ .
- 3)  $\text{Re}(G(jw)) > 0 \quad \forall w \geq 0$

- La première condition est vérifiée

- Pour vérifier la deuxième condition, appliquons le critère de ROUTH au dénominateur de  $G'(W)$  :

$$\begin{array}{l|l} W^2 & \begin{array}{l} 1 + a_1 - a_2 \\ 1 - a_1 - a_2 \end{array} \\ W^1 & \begin{array}{l} 2(1 + a_2) \\ 0 \end{array} \\ W^0 & \begin{array}{l} \frac{2(1 + a_2)(1 - a_1 - a_2)}{2(1 + a_2)} \\ 0 \end{array} \end{array}$$

Les conditions s'expriment donc:

$$(I) \quad \begin{cases} 1 + a_1 - a_2 > 0 \\ 2 + 2 a_2 > 0 \\ 1 - a_1 - a_2 > 0 \end{cases}$$

-Pour la 3ème condition, nous avons :

$$\Pi(w^2) = R_e(G(jw)) = R_e \left\{ \frac{(1 + c_1 + c_2) - (1 + c_2 - c_1)w^2 + (2 - 2c_2) jw}{(1 - a_1 - a_2) - (1 + a_1 - a_2)w^2 + (2 + 2a_2) jw} \right\} = \frac{N}{D}$$

En multipliant numérateur et dénominateur par le conjugué de D, on a :

$$\Pi(w^2) = \frac{1}{|D|^2} R_e \left\{ [(1 + c_1 + c_2) - (1 + c_2 - c_1)w^2 + (2 - 2c_2) jw] [(1 - a_1 - a_2) - (1 + a_1 - a_2)w^2 - (2 + 2a_2) jw] \right\}$$

$$\begin{aligned} \Pi(w^2) &= \frac{1}{|D|^2} \left[ (1 + c_1 + c_2)(1 - a_1 - a_2) - (1 + c_1 + c_2)(1 + a_1 - a_2)w^2 - (1 + c_2 - c_1) \cdot \right. \\ &\quad \left. \cdot (1 - a_1 - a_2)w^2 + (1 + c_2 - c_1)(1 + a_1 - a_2)w^4 + (2 - 2c_2)(2 + 2a_2)w^2 \right] \\ &= \frac{1}{|D|^2} \left\{ (1 + c_1 + c_2)(1 - a_1 - a_2) + [(2 - 2c_2)(2 + 2a_2) - (1 + c_1 + c_2)(1 + a_1 - a_2) \right. \\ &\quad \left. - (1 + c_2 - c_1)(1 - a_1 - a_2)] w^2 + (1 + c_2 - c_1)(1 + a_1 - a_2)w^4 \right\} \end{aligned}$$

Cette dernière condition exige donc que :

$$\Pi(w^2) > 0 ; \forall w \gg 0$$

Pour cela, il suffit que tous les coefficients de  $\Pi(w^2)$  soient non négatifs; par conséquent, on doit avoir :

$$(II) \left\{ \begin{array}{l} (1+c_1+c_2)(1-a_1-a_2) > 0 \\ (2-2c_2)(2+2a_2)-(1+c_1+c_2)(1+a_1-a_2)-(1+c_2-c_1)(1-a_1-a_2) > 0 \\ (1+c_2-c_1)(1+a_1-a_2) > 0 \end{array} \right.$$

D'après les conditions obtenues précédemment pour la stabilité de  $G'(p)$ , on déduit que :

$$(I) \text{ et } (II) \implies (III) \left\{ \begin{array}{l} 1+c_1+c_2 > 0 \\ (1+3a_2)-a_1c_1-c_2(3+a_2) > 0 \\ 1+c_2-c_1 > 0 \end{array} \right.$$

Les trois inéquations formulées dans le système (III) fournissent un ensemble de valeurs  $c_1$  et  $c_2$  du correcteur qui permettent de maintenir la condition de stricte réelle positivité de la partie linéaire, compte tenu de l'imprécision sur la connaissance des paramètres du procédé (lesquels sont déterminés par exemple à l'aide de la méthode des moindres carrés simples).

CHAPITRE VII

R E S U L T A T S    E X P E R I M E N T A U X

VII-1    DESCRIPTION DE LA MACHINE A PAPIER

=====

Le processus de fabrication du papier est très complexe et assez mal connu. Il est difficile d'exprimer mathématiquement les phénomènes physiques qui interviennent. La matière première est souvent hétérogène et il est difficile d'en définir les propriétés. Avec une même machine, on peut obtenir de nombreuses qualités de papier et pour chaque qualité trouver un grand nombre de paramètres de fonctionnement.

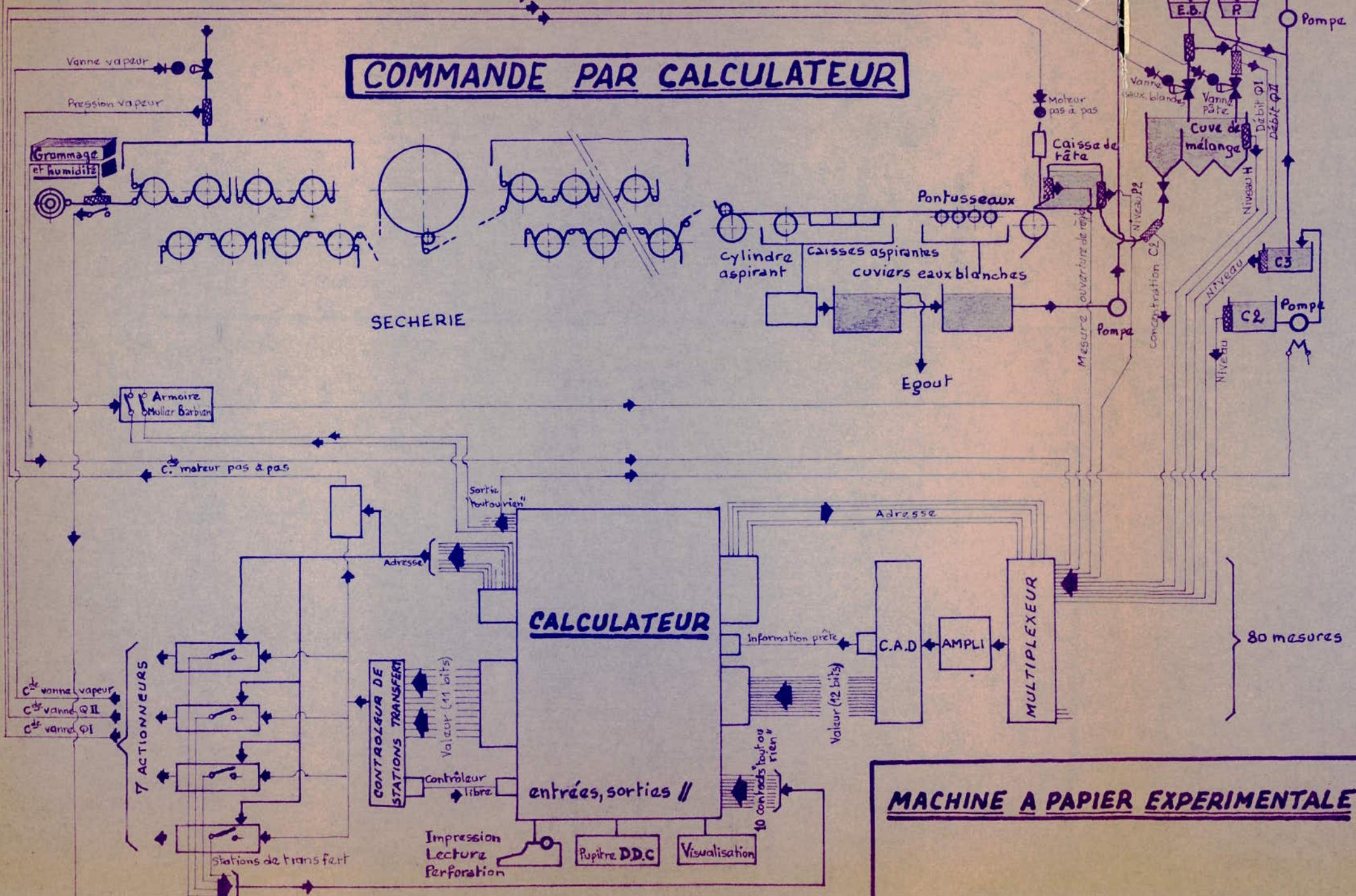
Parmi les nombreuses qualités du papier, on peut citer :

- Le grammage : poids par unité de surface,
- L'humidité : teneur en eau de la feuille en fin de chaîne de fabrication,
- L'épaisseur,
- La rigidité
- La blancheur ou la couleur, etc...

On distingue trois phases dans le processus qu'est la machine à papier :

- Phase de préparation de la pâte
- Phase de formation de la feuille: c'est la "partie humide" de la machine où se forme la feuille de papier
- Phase de séchage: à ce dernier stade, la feuille est séchée (au moyen de cylindres chauffés à la vapeur) et l'humidité finale (la teneur en eau) mesurée.

# COMMANDE PAR CALCULATEUR



## MACHINE A PAPIER EXPERIMENTALE

### Mesures sur le processus

- Débit de pâte : la mesure du débit de pâte s'effectue à l'aide de débitmètres électromagnétiques

- Grammage et humidité : la mesure est effectuée au moyen d'un capteur spécial (combiné Müller-Barbieri). Celui-ci donne la mesure de la masse au m<sup>2</sup> totale de la feuille par la mesure de l'absorption d'un rayonnement radioactif par le papier et la mesure de l'humidité par la mesure de la constante diélectrique de la feuille à l'aide d'un condensateur coplanaire.

- Détermination du grammage sec : à partir de la mesure du grammage total (matière sèche plus eau), et de la mesure de l'humidité, le calculateur détermine le grammage sec.

La figure de la page suivante représente le schéma de l'installation et du système de commande numérique. Le calculateur agit directement sur les actionneurs; les entrées-sorties du calculateur sont en parallèle. Son exploitation en temps réel nécessite un système d'interruption prioritaire. Pour améliorer le taux d'occupation de l'unité centrale, un système d'exploitation 'partagé' (temps partagé-temps réel) gère les tâches les moins prioritaires

### VII-2 TECHNIQUES DE MISE EN OEUVRE DES METHODES D'IDENTIFICATION =====

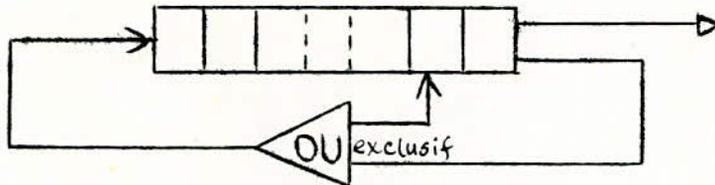
Avant d'exposer les techniques de mise en oeuvre de l'identification en temps réel, il est important de savoir que la notion de temps réel employée dans cette étude a un autre sens que celui employé en pratique. En effet, quand on parle de temps réel (TR), on suppose que le calculateur est directement couplé au processus qu'il commande ou qu'il observe. Dans notre cas, il s'agit d'une simulation du TR : nous avons emmagasiné les données en mémoire, et le programme utilisant un algorithme en TR les exploite de manière consécutive (pas à pas), exactement comme si on procédait à une acquisition

immédiate à partir d'un processus.

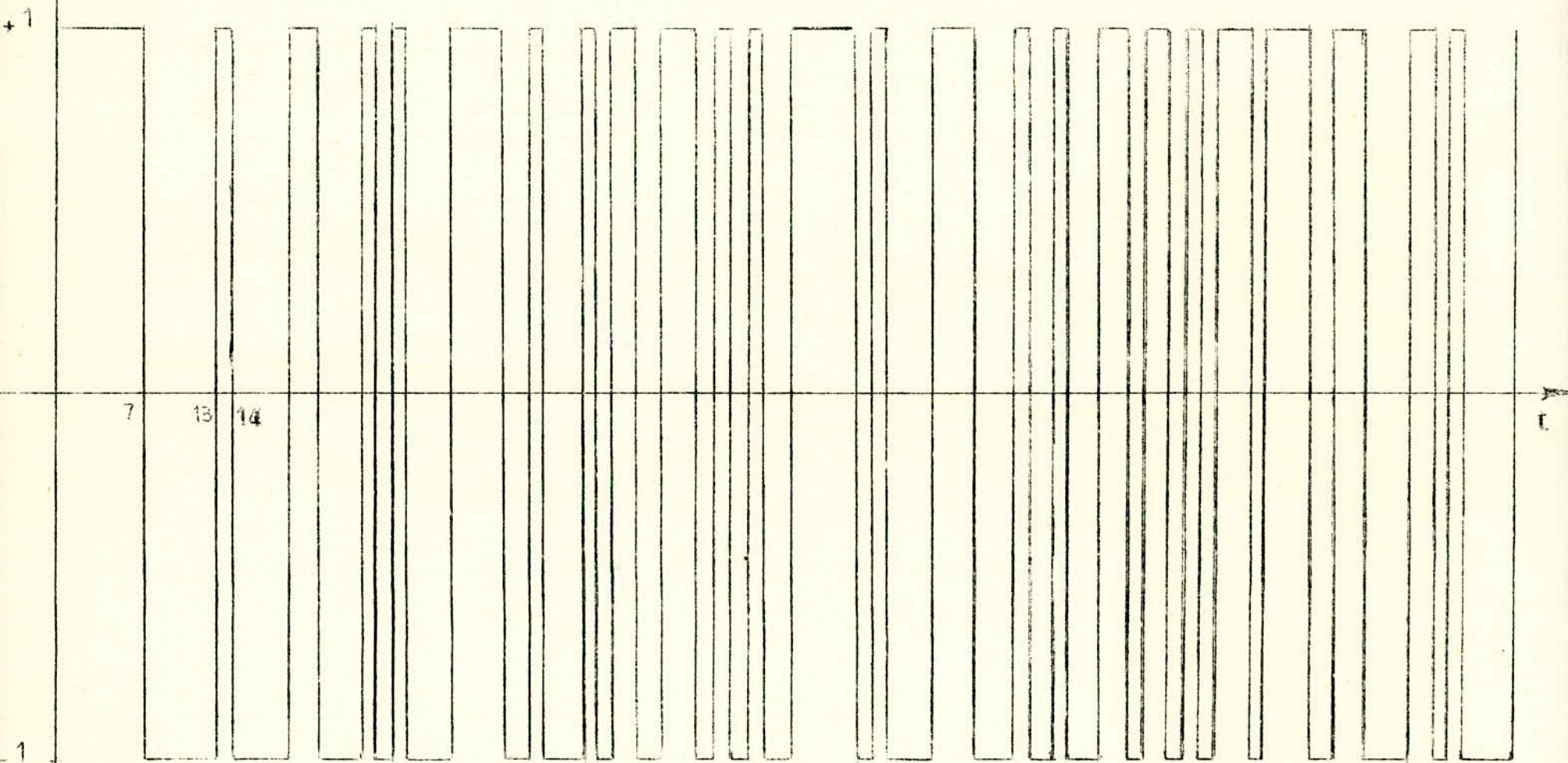
A ce propos, il convient de noter que la détermination du point de fonctionnement ne s'obtient pas par un moyennage sur l'ensemble des données emmagasinées, mais de manière progressive c'est-à-dire que ce point de fonctionnement est remis à jour à l'arrivée de chaque nouvelle donnée.

### VII-2-1 Choix de l'entrée

Le choix de l'entrée est d'une très grande importance pour l'identification. Le problème qui se pose est de pouvoir exciter tous les paramètres du procédé à identifier. Un bruit blanc serait le signal idéal pour cette opération, car comme on le sait, il est à spectre constant, donc à énergie égale pour chacune des fréquences qui le composent. L'expérience a montré aux automaticiens qu'un signal répondant à un tel besoin existe sous forme de séquence binaire pseudo-aléatoire (S.B.P.A.) que l'on prend de longueur maximale. Son spectre se rapproche alors de celui du bruit blanc; sa période peut être choisie égale à la plus petite des constantes de temps du système pour satisfaire au théorème de SHANNON. La génération d'une SBPA s'obtient en câblé grâce à un registre à décalage dont l'entrée est fournie par une porte OU<sub>ex</sub> à deux entrées. Ces deux entrées sont le dernier et le ième bits de ce même registre. Cette séquence peut être également obtenue par software grâce à une subroutine qui porte le nom de GAUSS-RANDU



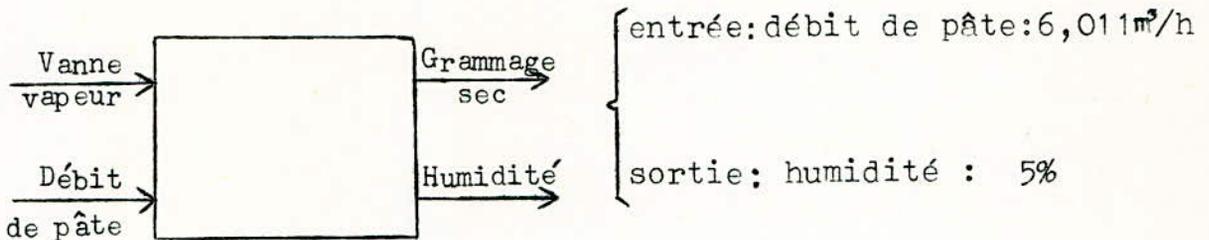
Dans notre cas, nous disposons d'un package conversationnel de 258 cartes: l'entrée du système variant autour d'un point de fonctionnement nominal, suivant le rythme d'une SBPA, l'enregis-



SEQUENCE BINAIRE PSEUDO-ALEATOIRE

trement représente donc le fonctionnement dynamique du système.

- Le point de fonctionnement nominal pour le transfert débit de pâte-humidité est le suivant:



- La période choisie lors d'études effectuées sur cette machine (Centre Technique de l'industrie des Papiers, Cartons et Celluloses de Gières, France) est de 5 secondes.

- L'amplitude de la SBFA est d'environ  $\pm 0,25 \text{ m}^3/\text{h}$ , car il faut tenir compte de l'erreur de mesure, de l'inertie des capteurs et des actionneurs.

Notons cependant que le point de fonctionnement ci-dessus n'est valable que pour l'horizon sur lequel nous avons opéré: 200 données soit  $200 \times 5 = 1000 \text{ s}$  ou environ 17 minutes de fonctionnement. Il peut en effet dériver et fausser l'identification. C'est pourquoi il est nécessaire parfois de le resituer sur des horizons allant de 5 à 10 périodes d'échantillonnage.

## VII-2-2 Initialisation des algorithmes et programmes de simulation

### a) Initialisation

Comme nous l'avons déjà noté, les algorithmes sont initialisés de façon adéquate au départ de toute identification. Leur forme générale peut être ramenée à la suivante:

$$\hat{\underline{\theta}}_{N+1} = \hat{\underline{\theta}}_N + \underline{K}_{N+1} \underline{\epsilon}$$

$$\underline{K}_{N+1} = f(\underline{P}_N, \underline{x}(N+1), \underline{x}^T(N+1))$$

$$\underline{P}_{N+1} = g(\underline{P}_N, \underline{K}_{N+1}, \underline{x}(N+1))$$

L'initialisation se fait donc comme suit:

$$\hat{e}_0 = \underline{0} \quad ; \quad P(I,J) = \delta_{i,j} \quad \begin{cases} \delta = 0 \text{ pour } i \neq j \\ \delta = 1000 \text{ pour } i=j \end{cases}$$

Pour l'estimation des résidus en TR, tous les paramètres  $\alpha$ ,  $\beta$ ,  $[\Gamma]$  sont initialisés à zéro.

Par ailleurs, nous savons que ces méthodes ne nécessitent pas de conditions initiales pour les entrées-sorties. Il faut toutefois tenir compte du décalage entre l'entrée et la sortie introduit par le retard pur: on perd ainsi  $R=13$  données, auxquelles viennent s'ajouter les  $n+p+1$  permettant d'écrire la première équation. Il convient en plus d'en sacrifier  $K$  pour la méthode de la matrice instrumentale,  $n$  pour les moindres carrés généralisés. Si on veut résumer ces pertes d'informations, on a :

-Moindres carrés	: $R+n+p+1$
-Matrice instrumentale	: $R+n+p+1+K$
-Moindres carrés généralisés	: $R+n+p+1+n$
-Hyperstabilité	: $R+n+p+1$

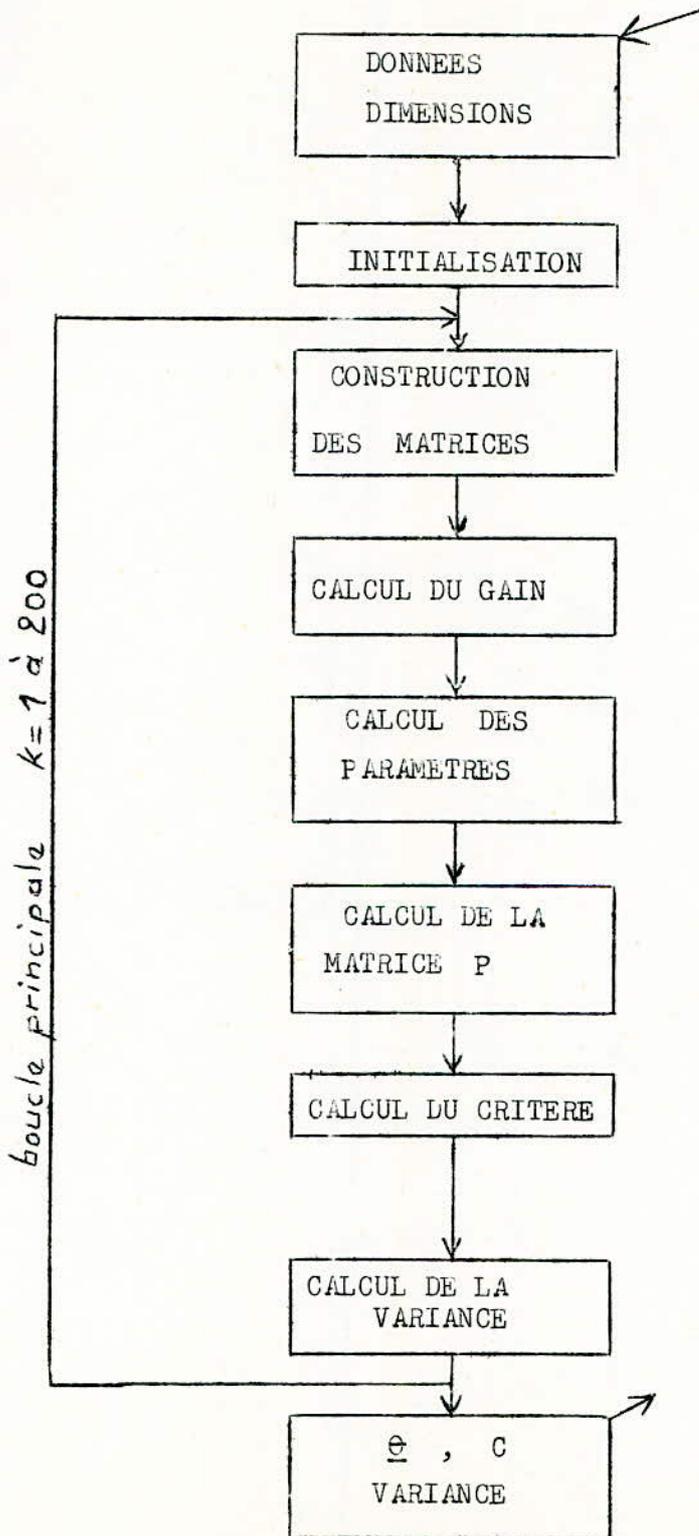
### b) Organigramme général

L'organigramme qui suit montre d'une manière simplifiée la décomposition de l'algorithme global (tronc commun aux différentes méthodes) lors de son traitement dans le programme de calcul. La boucle principale est celle du traitement complet pour chaque donnée, les boucles des calculs intermédiaires ont été omises pour plus de clarté.

### c) Programmes de simulation

Les programmes de simulation ont pour but la vérification d'un programme avant son application à une identification; c'est donc un moyen de vérifier si le programme est opérationnel et d'évaluer la précision sur les paramètres. Pour cela, on se donne un modèle, de paramètres connus ainsi que le retard pur et éventuellement la nature et le niveau d'un bruit. Il est

ORGANIGRAMME GENERAL

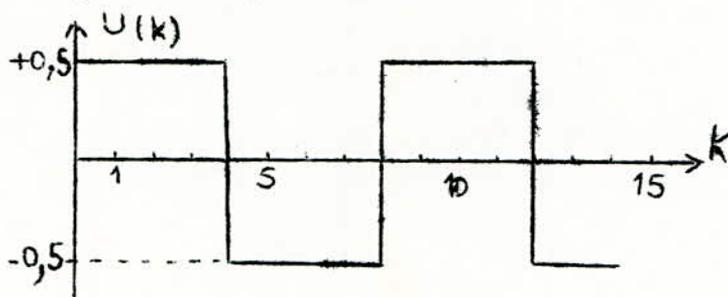


évident que le modèle de simulation doit être choisi stable, de façon à éviter la divergence des paramètres à estimer.

Nous avons retenu un modèle de simulation non bruité d'équation :

$$y(t) - 1,4 y(t-1) + 0,4 y(t-2) = 0,2 u(t-5) + 0,25 u(t-1-5)$$

Ce modèle est excité non pas par une SBPA (qui n'est pas nécessaire pour ce cas idéal), mais une séquence alternée de niveau  $\pm 0,5$  et de période 4



RESULTATS DE SIMULATION  
POUR LES DIFFERENTES METHODES ETUDIEES

	$a_1 = -1,4$	$a_2 = 0,4$	$b_0 = 0,2$	$b_1 = 0,25$	Critère	Précision
	$\hat{a}_1$	$\hat{a}_2$	$\hat{b}_c$	$\hat{b}_1$		
M.C	-1,3996	0,3996	0,1999	0,2501	0,9740	$4 \cdot 10^{-4}$
M.I	-1,3985	0,3996	0,1992	0,2501	2,4921	$1,5 \cdot 10^{-3}$
M.C.G	-1,3996	0,3996	0,1999	0,2501	0,9585	$4 \cdot 10^{-4}$
HYP *	-1,403	0,403	0,203	0,247	/	$3 \cdot 10^{-3}$

-\* Pour la simulation de cette méthode nous avons utilisé un correcteur parfait, c'est-à-dire :

$$c_1 = a_1 \quad \text{et} \quad c_2 = a_2$$

VII-2-3 Détermination du retard pur de la machine  
à papier

Nous avons vu au chapitre II que le retard pur étant une constante de la machine et qu'il représente un certain nombre de petites constantes de temps de cette machine. On peut le caractériser indifféremment par l'une ou l'autre des méthodes. Le principe de sa détermination est le suivant : considérons le critère  $C = \sum_{t=1}^R e^2(t)$ , somme des carrés des erreurs de prédiction. En faisant varier R de 0 à 20 par exemple, nous obtenons une courbe passant par un minimum. Cette valeur donnant l'erreur minimale est donc la plus convenable. Nous avons opéré avec la méthode des moindres carrés simples; on obtient:

R	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
C	9,26	9,03	6,58	6,48	8,21	8,22	10,43	2,39	2,55	2,55	2,76	2,08	1,81	1,56	2,06	2,55	3,86	3,91

Le critère minimal s'obtient pour  $R = 13$ ; cela signifie que le retard pur global dans le transfert débit pâte-humidité est  $13 \times 5 = 65$  s. Ainsi, la sortie n'est sensible à une variation de l'entrée que 65s plus tard; ceci dû, rappelons-le au transport de matières, à l'inertie des actionneurs (moteurs pas à pas),...

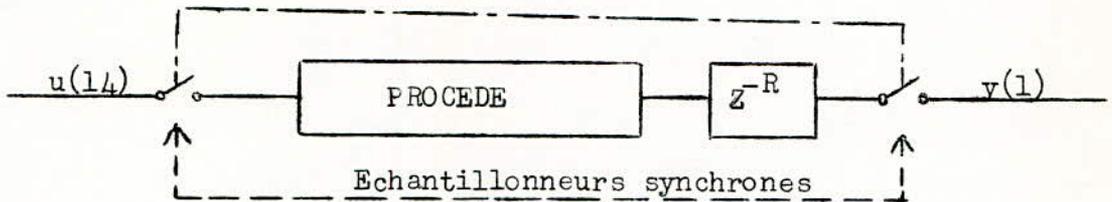
REMARQUE :

===== a) Il est vrai que le fait de choisir un nombre entier de périodes (choix imposé par le type de modèle aux différences) introduit une certaine erreur, mais on conçoit qu'elle peut être négligée devant les erreurs à l'acquisition des mesures et à la transmission des ordres d'action.

b) L'ordre du modèle, qui n'est pas encore déterminé, n'influe pas sur la détermination de R qui est "intrinsèque" au processus. Nous avons pris  $n = 2$ , valeur la plus probable

(pour cela cf B1).

c) Le retard pur se traduit à l'enregistrement par une translation de R périodes des sorties par rapport aux entrées. En effet, à la prise des données sur le processus, l'échantillonnage est synchrone



On enregistre donc simultanément  $u(14)$  et  $y(1)$ ;  $u(15)$  et  $y(2)$  etc...

Le retard étant choisi, on a un modèle de la forme:

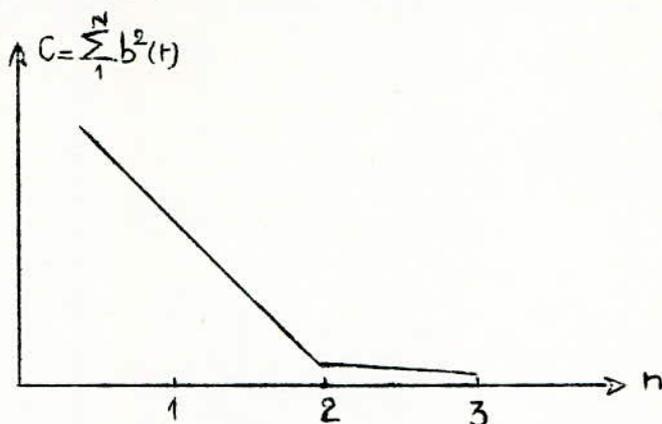
$$\begin{aligned} y(t-13) + a_1 y(t-14) + \dots + a_n y(t-13-n) &= \\ &= b_0 u(t) + b_1 u(t-1) + \dots + b_p u(t-p) \end{aligned}$$

#### VII-2-4 Caractérisation de l'ordre du modèle

Le choix de l'ordre du modèle ou caractérisation, peut se faire, comme le retard pur, indifféremment par l'une ou l'autre des méthodes étudiées. Cette caractérisation est primordiale car un mauvais choix de l'ordre risque de déboucher sur un modèle tout à fait éloigné de la représentation optimale que l'on recherche.

Le principe de la détermination est la même que pour celle du retard pur: on se base sur l'évolution du critère lorsque l'on fait varier  $n$ . On obtient dans tous les cas, d'une manière d'ailleurs plus ou moins probante, une courbe dite courbe de caractérisation  $C=f(n)$ ; celle-ci, si la caractérisation est bonne, présente une première cassure nette pour l'ordre optimal du modèle. D'autres cassures suivent, plus atténuées pour les

ordres supérieurs (parfois, on constate une nouvelle croissance du critère)



Nous avons opéré dans notre cas, avec la méthode des moindres carrés généralisés avec 200 données. Les résultats obtenus sont consignés dans le tableau 1.

Le plus petit critère est 1,479, correspondant à l'ordre 2: nous en déduisons que le transfert débit pâte-humidité est assimilable à un système linéaire du second ordre. Nous retiendrons cette valeur pour toute la suite.

### VII-3 RESULTATS D'IDENTIFICATION DES PARAMETRES DU TRANSFERT ===== DEBIT PATE-HUMIDITE DE LA MACHINE A PAPIER =====

Nous allons présenter les résultats d'identification des paramètres sous forme de tableaux à deux ouvertures: l'ordre étant choisi égal à 2, la complexité du modèle est fonction de  $p$ . Comme on doit avoir  $p \leq n=2$ , on a trois possibilités:

- $p = 0$
- $p = 1$
- et -  $p = 2$

En plus des paramètres  $\theta_i$ , il est utile de consigner la valeur du critère pour la comparaison des trois modèles.

TABLEAU 1

RESULTATS DE CARACTERISATION DE L'ORDRE

Méthode : Moindres carrés généralisés

Critère prédictif :  $C = \sum_{k=1}^N b^2(t)$

$p = 1$

n	1	2	3
C	1,558	1,479	1,578

REMARQUE : le paramètre p qui est l'ordre du numérateur de la fonction de transfert recherchée :

$$F(Z^{-1}) = \frac{B(Z^{-1}) u(t)}{A(Z^{-1}) y(t)} = \frac{\sum_{i=0}^p u(t-i) Z^{-i}}{1 + \sum_{i=1}^n y(t-i) Z^{-i}}$$

représente la complexité du modèle. Il n'influe pas sur la caractérisation de n. Nous l'avons pris égal à 1 pour "uniformiser" cette caractérisation. Le modèle le plus complexe pour n = 1 étant p = 1 et pour n = 2; 3 ; p = 2; 3 il est tout à fait logique de prendre p = 1 pour pouvoir comparer les critères avec des modèles de même complexité et d'en déduire la courbe de caractérisation.

TABLEAU 2

METHODE DES MOINDRES CARRES SIMPLES  
 RESULTATS D'IDENTIFICATION AVEC 200 DONNEES

$$\text{Critère prédictif } C = \sum_{t=t_0+1}^N e^2(t)$$

n = 2
-------

$\theta_i \backslash p$	0	1	2
$a_1$	-1,648	-1,535	-1,482
$a_2$	0,688	0,580	0,531
$b_0$	0,174	0,155	0,155
$b_1$	—	0,128	0,129
$b_2$	—	—	0,036
C	1,686	1,470	1,263

\* Le critère minimal s'obtient donc pour le modèle le plus complexe :  $p = 2$  .

TABLEAU 3

MATRICE INSTRUMENTALE  
CARACTERISATION DU RETARD K - 200 DONNEES -

$$\text{Critère prédictif } C = \sum_{n=1}^N e^2(t)$$

$n = 2$ , $p = 2$
-------------------

K	2	3	4
C	9,988	1,888	187

\* La détermination de K ,retard des observations, s'effectue de la même manière que celle du retard pur: on fait varier K et on opte pour la valeur donnant un critère minimal , pour un modèle fixe ( n=2 , p=2).

TABLEAU 4

METHODE DE LA MATRICE INSTRUMENTALE  
 RESULTATS D'IDENTIFICATION AVEC 200 DONNEES

$$\text{Critère prédictif } C = \sum_{n+1}^N e^2(t)$$

$n = 2$ , $K = 2$
-------------------

$\theta_i \backslash p$	0	1	2
$a_1$	-1,340	-1,400	-1,420
$a_2$	0,391	0,447	0,470
$b_0$	0,170	0,156	0,154
$b_1$	—	0,146	0,138
$b_2$	—	—	0,053
$C$	28,4	2,15	9,99

\* Dans la partie théorique , nous avons vu que BANON a proposé  $K = n$  : les résultats obtenus avec cette valeur peuvent être comparés à ceux obtenus avec  $K = 3$  , valeur optimale déterminée par caractérisation ( cf tableau 5 )

TABLEAU 5

METHODE DE LA MATRICE INSTRUMENTALE  
 RESULTATS D'IDENTIFICATION AVEC 200 DONNEES

Critère prédictif  $C = \sum_{t=1}^N e^2(t)$

$n = 2$  ,  $K = 3$

$e_i \backslash p$	0	1	2
$a_1$	-1,074	-1,271	-1,400
$a_2$	0,143	0,330	0,451
$b_0$	0,167	0,148	0,153
$b_1$	—	0,172	0,144
$b_2$	—	—	0,06
C	14,60	> 1000	1,888

TABLEAU 6

METHODE DES MOINDRES CARRES GENERALISES  
 RESULTATS D'IDENTIFICATION AVEC 200 DONNEES

$$\text{Critère prédictif } C = \sum_{t=1}^N b^2(t)$$

$$n = 2$$

$\theta_i \backslash p$	0	1	2
$a_1$	-1,730	-1,574	-1,479
$a_2$	0,764	0,615	0,527
$b_0$	0,178	0,158	0,155
$b_1$	—	0,123	0,129
$b_2$	—	—	0,035
C	1,672	1,479	1,318

TABLEAU 7

METHODE DE L'HYPHERSTABILITE  
 RESULTATS D'IDENTIFICATION AVEC 200 DONNEES

Critère prédictif  $C = \sum_{n+1}^N \mathcal{E}^2(t)$

$n = 2$

$\theta_i \backslash p$	0	1	2
$a_1$	-1,802	-1,607	-1,550
$a_2$	0,828	0,643	0,592
$b_0$	0,170	0,156	0,156
$b_1$	—	0,114	0,122
$b_2$	—	—	0,021
$C$	18,06	16,38	18,09

VII-4 LES MODELES OBTENUS ET LEURS PERFORMANCES

=====

Pour chacune des méthodes, nous sommes en possession de trois modèles possibles. Nous nous sommes basés ici sur le critère de prédiction pour retenir le plus précis; il s'avère que dans chaque cas, hormis l'hyperstabilité, le critère minimal s'obtient pour le modèle le plus complexe, c'est-à-dire pour  $n = p = 2$ .

Ce choix est motivé ici par le but suivant : on se propose de vérifier au moyen de courbes, la poursuite du procédé par le modèle. Il est entendu que lors de l'application à une commande directe, il faudra tenir compte du temps de calcul et du coût pour remodeler ce choix, qui sera le résultat d'un compromis entre la précision et la simplicité.

Donc pour cette étude, nous avons retenu les modèles suivants:

M.C	$y(t)-1,482y(t-1)+0,531y(t-2)=0,155u(t-13)+0,129u(t-14)+0,036u(t-15)$
M.I	$y(t)-1,4y(t-1)+0,451y(t-2)=0,153u(t-13)+0,144u(t-14)+0,06u(t-15)$
M.C.G	$y(t)-1,479y(t-1)+0,527y(t-2)=0,155u(t-13)+0,129u(t-14)+0,035u(t-15)$
HYP	$y(t)-1,55y(t-1)+0,592y(t-2)=0,156u(t-13)+0,122u(t-14)+0,021u(t-15)$

Pour visualiser la poursuite modèle-procédé dans chaque cas, nous avons tracé les courbes:

$$y_p = f(u, k)$$

$$y_m = g(u, k)$$

L'évaluation de l'écart entre sortie du procédé et sortie du modèle, est faite dans chaque cas à l'aide d'un critère quadratique moyen, dont l'expression est la suivante:

$$J = \frac{1}{N} \cdot \sum_{K=1}^N (y_p(K) - y_m(K))^2$$

Ce critère donne un ordre de grandeur de l'erreur moyenne et permet d'établir la comparaison entre les performances des modèles retenus

Méthode	Critère J
M.C	0,0852
M.I	0,0928
M.C.G	0,0879
HYP	0,0847

### Analyse des courbes obtenues

Nous étions en possession de 258 couples de valeurs entrées-sorties. Nous avons mené l'identification avec les 200 premières. La comparaison sortie du modèle-sortie du procédé a utilisé l'ensemble des 258 points, ceci dans le but de voir le comportement du modèle dans l'ensemble des points avec lesquels l'identification a été menée d'une part, et d'autre part les 58 derniers, qui ne sont pas entrés en jeu dans cette identification;

Sur toutes les courbes, on remarque que la poursuite est satisfaisante, surtout entre le 160e et le 240e points. L'écart est sensiblement plus grand au départ: ceci est normal, puisque les paramètres du modèle retenu pour chaque méthode choisis dans l'intervalle correspondant à k [160-240]

REMARQUE

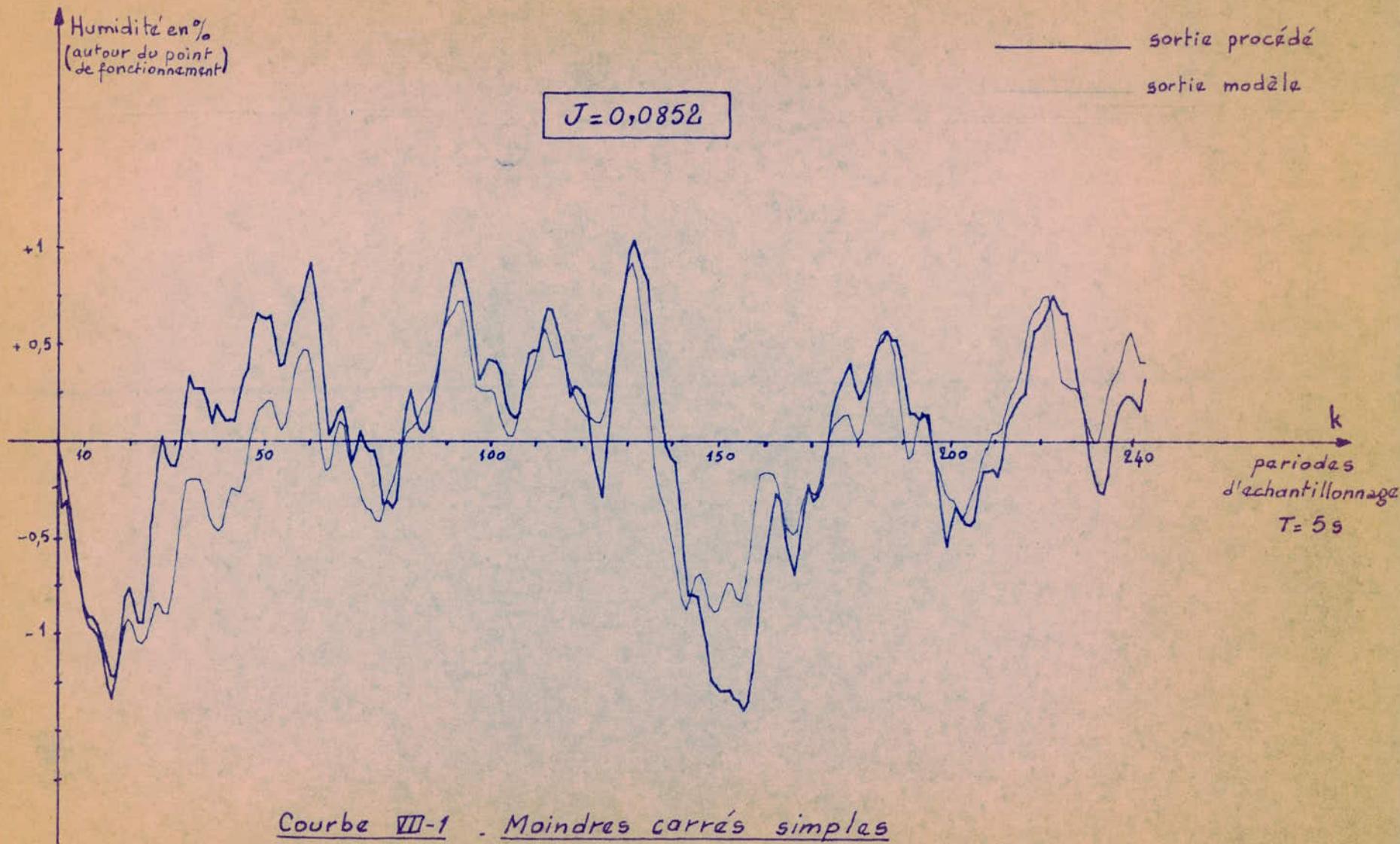
=====

Ces courbes donnent seulement un aperçu de l'évolution des sorties et par conséquent, de l'écart. D'une façon très précise, la sortie du modèle doit être calculée à chaque instant avec les paramètres obtenus à cet instant, car ceux-ci -du fait de la présence du bruit- ne sont pas figés, mais plus ou moins oscillants, même en fin d'identification. Une façon commode d'améliorer cette poursuite est de prendre des paramètres moyennés sur des horizons de 5 valeurs par exemple (modèle "à paramètres variables")

\*\*\* Le tracé de ces courbes a été effectué au moyen d'une calculatrice programmable HEWLETT-PACKARD munie d'une table traçante

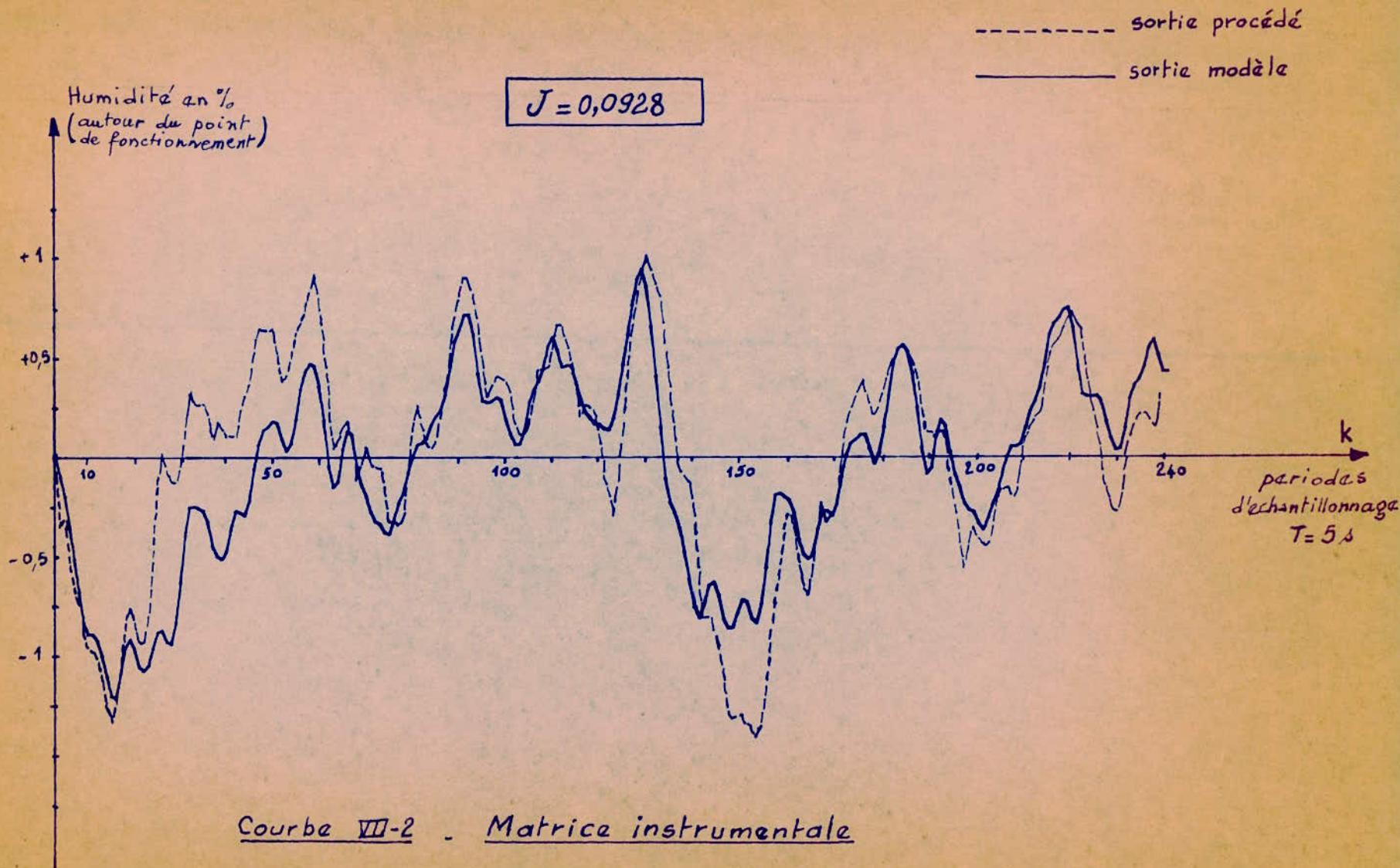
Cet appareil, de même que le calculateur IBM 1130 sur lequel nous avons effectué nos calculs se trouvent au CENTRE de CALCUL de l'Ecole.

-----oOo-----



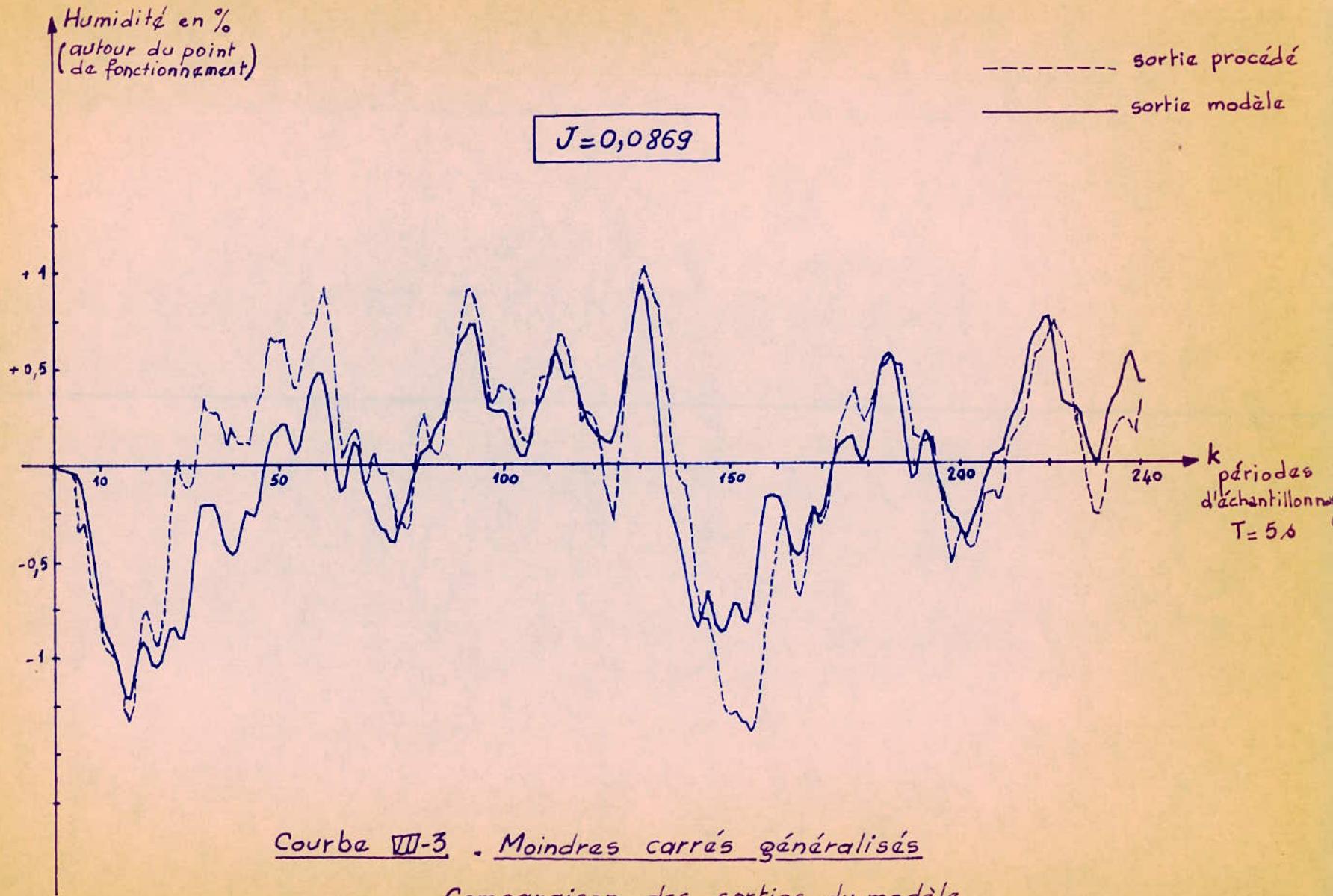
Courbe VII-1 . Moindres carrés simples

Comparaison des sorties du modèle  
et du procédé



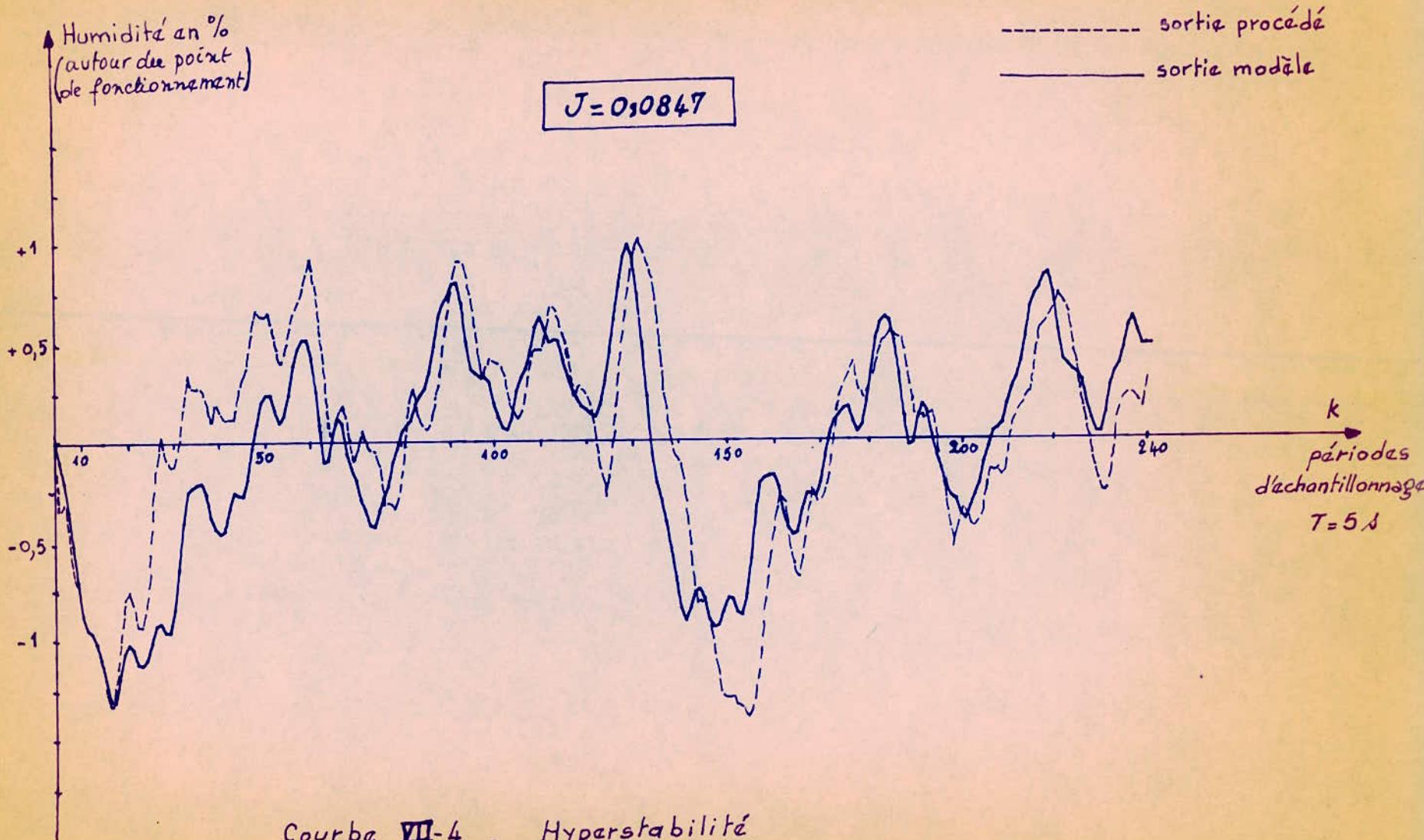
Courbe VII-2 - Matrice instrumentale

Comparaison des sorties du modèle  
et du procédé



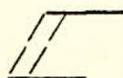
Courbe VII-3 . Moindres carrés généralisés

Comparaison des sorties du modèle  
et du procédé



Courbe VII-4 - Hyperstabilité

Comparaison des sorties du modèle  
et du procédé



## CONCLUSION

Cette étude nous a permis de déterminer la fonction de transfert débit pâte-humidité de la machine à papier expérimentale. Les méthodes utilisées peuvent être appliquées à d'autres transferts dans cette même machine, comme par exemple le transfert débit pâte-grammage sec avec une procédure tout à fait identique à celle adoptée dans ce projet. D'une manière plus générale, ces méthodes peuvent être adaptées à d'autres procédés industriels, ce qui se fait d'ailleurs actuellement, de façon de plus en plus courante. Les résultats d'identification permettent par la suite d'élaborer des systèmes de commande qui assurent de meilleures performances dans le fonctionnement de ces procédés (amélioration de la qualité des produits et du rendement, de la sécurité).

Les méthodes que nous avons présentées diffèrent quelque peu les unes des autres sur le plan de la précision et sur le plan de la complexité:

- Les moindres carrés simples sont les plus explicites et les plus rapides; avec cette méthode, nous avons obtenu de bons résultats, comparables à ceux des autres méthodes qui sont pourtant plus élaborées. Ce fait peut s'expliquer par la présence d'un niveau de bruit faible. Dans le cas d'un bruit de niveau élevé, cette méthode donne des paramètres fortement biaisés et devient peu efficace

- La méthode de la matrice instrumentale est un peu plus complexe, mais permet d'éviter le biais de l'estimateur dans le cas de procédés très bruités

- Les moindres carrés généralisés à modèle élargi nécessitent des calculs assez longs, mais offrent l'avantage d'une meilleure précision grâce à la correction de l'erreur introduite dans la structure du modèle et aussi par la nature des résidus. Ces derniers sont les valeurs du bruit de sortie à chaque instant et sont donc moins corrélés avec les sorties (dans

la matrice  $X$ ) et le biais de l'estimateur est plus faible.

- La méthode de l'hyperstabilité nécessite elle, une identification préalable pour la détermination du correcteur linéaire, ce qui est un inconvénient. Plus complexe que celle des moindres carrés, la convergence des paramètres dans cette méthode est assez lente, mais présente l'avantage d'être sans oscillation : la convergence est asymptotique et pratiquement sans biais.

IL est difficile de choisir parmi ces méthodes, une qui réponde à toutes les exigences de l'utilisateur. Tout dépend de l'importance de l'installation, des performances du calculateur, de la précision recherchée et du prix que l'on veut consacrer à l'identification, puis éventuellement à la commande du processus. Il importe en effet d'établir au préalable le coût des dispositifs et celui de leur fonctionnement. Quoiqu'il puisse être élevé au départ, les améliorations apportées par la commande automatique directe à l'aide de mini-calculateurs sont censées compenser ce coût en quelques années d'utilisation puis devenir rentables

L'application à la commande automatique des techniques de multiprogrammation en temps réel permettent l'automatisation d'ensembles beaucoup plus complexes que celui que nous avons pris ici en exemple et le niveau d'efficacité atteint beaucoup plus important. On s'attache de plus en plus à l'emploi de plusieurs minicalculateurs couplés à différents systèmes d'une unité de production afin d'éviter un arrêt total, ce qui arrive dans le cas de gros calculateurs gérant à eux seuls l'ensemble des tâches.

B I B L I O G R A P H I E

- B1 - S. GENTIL : Etude comparative de diverses méthodes <sup>Statistiques</sup> d'identification de systèmes dynamiques. Thèse de 3ème cycle; Grenoble, 1972.
- B2 - B. LEBEAU : Commande par ordinateur numérique d'une unité pilote de fabrication de papier. Thèse de docteur-ingénieur ; Grenoble , 1971 .
- B3 - Y. SEVELY : Systèmes et asservissements linéaires échantillonnés . DUNOD , 1968 .
- B4 - R. BOUDAREL, P. GUICHET, J. DELMAS : Commande optimale des processus, tome I . DUNOD , 1967 .
- B5 - J.C. RADIX : Introduction au filtrage numérique EYROLLES , 1970 .
- B6 - I.D. LANDAU : Sur une méthode de synthèse des systèmes adaptatifs avec modèle utilisés pour la commande et l'identification d'une classe de procédés physiques. Thèse de doctorat ès sciences ; Grenoble , 1973 .