

# Diagnostic et observabilité des systèmes non linéaires

---

<b>2.1. Introduction</b> .....	9
<b>2.2. Diagnostic des systèmes dynamique</b> .....	9
2.2.1. Différents types de défaut.....	9
a. <i>Défauts capteurs</i> .....	10
b. <i>Défauts actionneurs</i> .....	11
c. <i>Défauts composants</i> .....	12
2.2.2. Classification des méthodes de diagnostic.....	13
a. <i>Méthodes sans modèle</i> .....	14
b. <i>Méthodes à base de modèle</i> .....	17
c. <i>Analyses comparatives des méthodes de diagnostic</i> .....	19
2.2.3. Détection, localisation, et identification des défauts à base d'observateur non linéaire.....	20
a. <i>Génération des résidus</i> .....	24
b. <i>La détection</i> .....	29
c. <i>La localisation</i> .....	31
d. <i>L'identification</i> .....	34
2.2.4. Caractéristiques souhaitables d'un système de diagnostic.....	35
<b>2.3. Observabilité et observateurs grand gain</b> .....	36
2.3.1. Généralité sur les observateurs.....	37
2.3.2. Système en considération.....	37
2.3.3. Observabilité des systèmes non linéaires.....	38
2.3.4. Formes normales d'observabilité.....	41
a. <i>Cas où <math>n_y &gt; n_u</math></i> .....	41
b. <i>Cas où <math>n_y \leq n_u</math></i> .....	42
2.3.5. Formes normales Multi-entrées Mono-sortie.....	44
2.3.6. Observateur à grand gain.....	45
a. <i>Observateur à grand gain de type Luenberger</i> .....	45
b. <i>Observateur de Kalman étendu grand gain</i> .....	46
2.3.7. Observateurs à grand gain adaptatifs.....	47
a. <i>Adaptation du gain pour les observateurs de type Luenberger</i> .....	47
b. <i>Filtre de Kalman adaptatif</i> .....	49
c. <i>Observateur proposé par Busvelle et Gauthier</i> .....	50
<b>2.4. Conclusion</b> .....	52

---

## 2.1. Introduction

Ce chapitre a essentiellement pour objectif de présenter quelques rappels sur des concepts fondamentaux relatifs au diagnostic et à l'observabilité des systèmes dynamiques non linéaires.

La première partie propose, tout d'abord, un état de l'art des différents types de défauts et de leurs influences sur le processus à diagnostiquer. Dans un deuxième volet, les différentes méthodes de diagnostic seront présentées, elles seront regroupées en deux grandes familles: les méthodes utilisant des modèles mathématiques et celles ne les utilisant pas. Un intérêt particulier sera accordé aux méthodes de détection et d'isolation des défauts à base d'estimateur d'états. Une étude des méthodes de génération de résidus à base d'observateurs non-linéaires sera abordée. Par la suite, une description des trois autres étapes d'un module de diagnostic: la détection, la localisation et l'identification des défauts, sera exposée en détail. Nous discuterons à la fin de cette partie des performances attendues d'une procédure de diagnostic.

Le travail présenté dans la seconde partie de ce chapitre s'inscrit dans le cadre général de la théorie d'observation développée par Gauthier et Kupka [Gauthier et Kupka, 2001]. Un résultat important de cette théorie montre qu'il existe des représentations des systèmes non linéaires qui caractérisent l'observabilité (nommées forme normale d'observabilité). Ainsi, nous commencerons par un rappel des notions fondamentales d'observabilité. Ensuite, nous aborderons les différents types d'observateurs à grand-gain: ceux de type de Luenberger et ceux de type de Kalman. Nous terminerons par une étude bibliographique non exhaustive des observateurs à grand gain adaptatifs.

## 2.2. Diagnostic des systèmes dynamiques

le diagnostic est un système d'aide à la décision qui permet de localiser les composants ou les organes défaillants d'un système et éventuellement de déterminer ses causes. En parcourant la littérature, on se rend compte immédiatement que la terminologie dans ce domaine n'est pas cohérente. De nombreuses définitions d'un même mot sont trouvées. Afin d'enlever toutes ambiguïtés, le comité technique SAFEPROCESS de l'IFAC (International Federation of Automatic Control) ont essayé de standardiser cette terminologie. Un rappel de ces définitions est données dans l'annexe D.

### 2.2.1. Différents type de défaut

Un défaut est défini comme étant un écart non permis entre la valeur réelle d'une caractéristique du système et sa valeur nominale. Comme le montre la figure (2.1), trois types de défauts sont distingués: défaut actionneur, défaut capteur et défaut procédé (ou défaut composant).

Chacun de ces défauts ainsi que leurs influences sur le processus sont décrits brièvement dans le paragraphe ci-dessous.

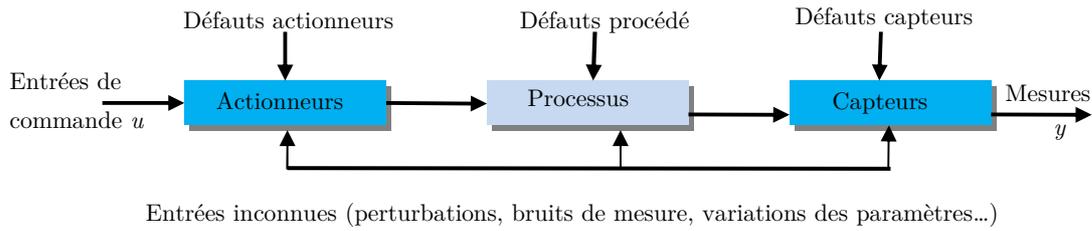


Figure 2.1. Différents types de défauts d'un système physique.

### a. Défauts capteurs

Les capteurs sont des instruments qui transforment une grandeur physique en une grandeur traitable par les calculateurs. Les capteurs sont essentiellement les interfaces de sortie d'un système avec l'environnement extérieur. Ils permettent de communiquer les informations concernant l'état et le comportement interne du processus. Ainsi, un défaut capteur caractérise une mauvaise image de la grandeur physique à mesurer. Pour les systèmes en boucle fermée, les mesures issues de ces capteurs sont utilisées pour la génération du signal de commande. Par conséquent, la présence d'un défaut capteur donne un signal de commande inexacte et inefficace.

Les défauts capteurs les plus répandus sont: a) le biais, b) la dérive, c) la perte d'efficacité, d) le blocage et e) le défaut de calibrage. La figure (2.2) représente l'effet de ces défauts sur les mesures.

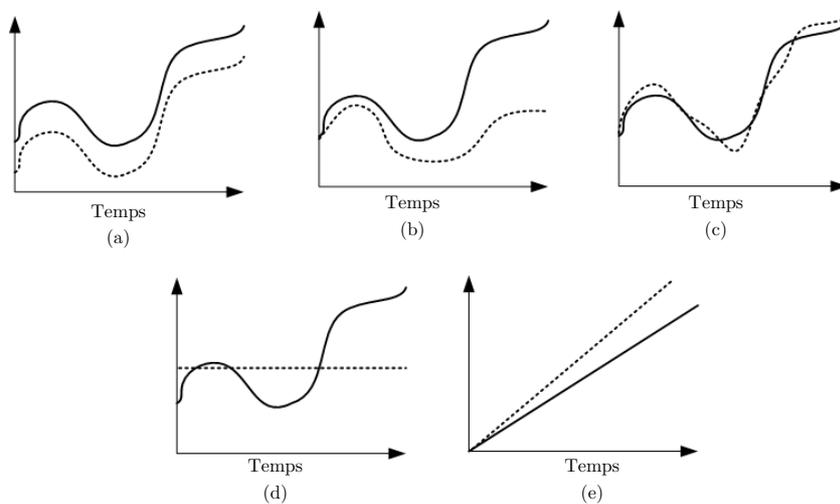


Figure 2.2. L'effet de différents types de défauts capteur sur les mesures. Les lignes en pointillées désignent les valeurs mesurées de capteur, cependant les lignes en trait continu représentent les valeurs réelles.

Les équations mathématiques de ces défauts sont comme suit [Sobhani-Tehrani et Khorasani, 2009]

$$y_i(t) = \begin{cases} x_i(t) & \forall t \geq t_0 & \text{sans défaut} \\ x_i(t) + b_i & b_i(t) = 0, b_i(t_{f_i}) \neq 0 & \text{biais} \\ x_i(t) + b_i(t) & |b_i(t)| = c_i(t), 0 < c_i \ll 1 & \forall t \geq t_{f_i} & \text{dérive} \\ x_i(t) + b_i(t) & |b_i(t)| \leq \bar{b}_i, \dot{b}_i(t) \in L^\infty & \forall t \geq t_{f_i} & \text{perte d'efficacité} \\ x_i(t_{f_i}) & & \forall t \geq t_{f_i} & \text{blocage} \\ k_i(t)x_i & 0 < \bar{k} \leq k_i(t) \leq 1 & \forall t \geq t_{f_i} & \text{défaut de calibrage} \end{cases}$$

Avec  $t_{f_i}$  c'est le temps d'occurrence d'un défaut pour l' $i^{\text{ème}}$  capteur et  $b_i$  désigne sa précision où  $b_i \in [-\bar{b}_i, \bar{b}_i]$  avec  $\bar{b}_i > 0$ . En plus,  $k_i \in [\bar{k}_i, 1]$  avec  $\bar{k}_i > 0$  représente l'efficacité minimale du capteur.

### b. Défauts actionneurs

L'actionneur est un élément de la partie opérative d'un système qui transforme les signaux de commande issus du contrôleur (micro-processeur, ou microcontrôleur) en mouvement (moteur), en chaleur (résistance électrique), ou en champ magnétique (électroaimant)...

Ainsi, les défauts actionneurs agissent au niveau de la partie opérative et détruisent le signal d'entrée du système. Les conséquences de défauts actionneurs peuvent varier d'une consommation élevée de l'énergie jusqu'à la perte totale de contrôle. Les défauts actionneurs varient d'un actionneur à un autre, mais une classification non exhaustive des défauts les plus fréquents est donnée figure (2.3).

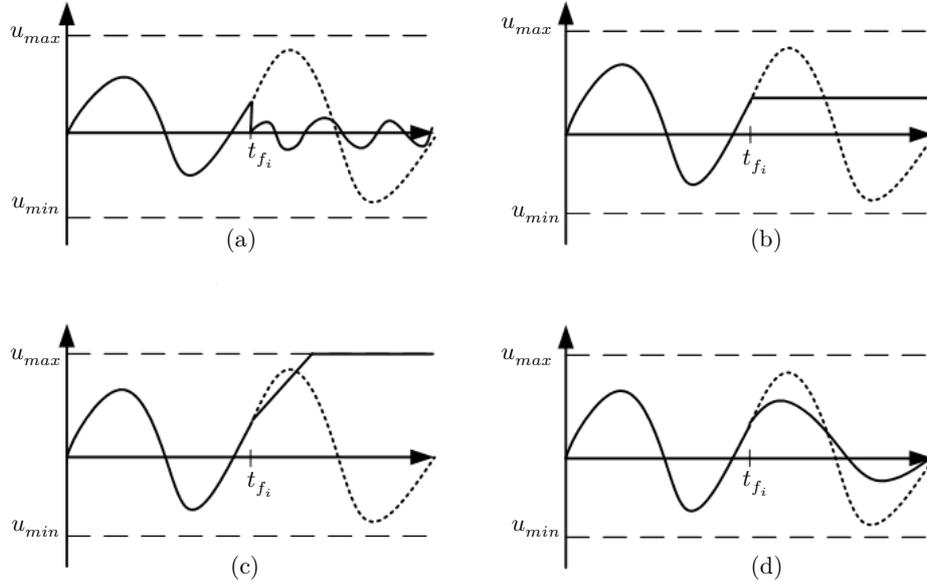


Figure 2.3. Représentation graphique des défauts actionneur les plus répandues [Ducard, 2009]. Les lignes en pointillées désignent les valeurs désirées de l'actionneur, cependant les lignes en trait continu représentent les valeurs actuelles. (a) oscillation, (b) blocage, (c) saturation et (d) perte d'efficacité.

Les différents types de défauts actionneurs peuvent être représentés par les expressions suivantes:

$$u_a^i(t) = \begin{cases} u_c^i(t) & 0 < \varepsilon_i \leq k_i(t) < 1, \forall t \geq t_{f_i} & \text{sans défaut} \\ k_i(t)u_c^i(t) & \forall t \geq t_{f_i} & \text{perte d'efficacité} \\ 0 & \forall t \geq t_{f_i} & \text{oscillation} \\ u_c^i(t_{f_i}) & \forall t \geq t_{f_i} & \text{blocage} \\ u_{i \min} \vee u_{i \max} & \forall t \geq t_{f_i} & \text{saturation} \end{cases}$$

Avec

$u_a^i(t)$  est le signal de sortie de l'actionneur ;  $u_c^i(t)$  est le signal de commande appliqué à l' $i^{\text{ème}}$  actionneur ;  $t_{f_i}$  désigne le temps d'occurrence d'un défaut à l' $i^{\text{ème}}$  actionneur ;  $k_i$  est l'efficacité de l'actionneur,  $\varepsilon_i > 0$  est l'efficacité minimale;  $u_{i \max}$  et  $u_{i \min}$  sont respectivement les limites supérieures et inférieures du signal de sortie de l' $i^{\text{ème}}$  actionneur.

### c. Défauts composants

Les défauts composants sont des défauts qui affectent les composants du système lui même. Ce sont les défauts qui ne peuvent pas être classifiés ni parmi les défauts actionneurs ni parmi les défauts capteurs. Ce type de défaut provoque un changement de la dynamique du système suite à un changement de ces paramètres. La représentation mathématique des défauts composants est souvent difficile à déterminer et demande des essais expérimentaux extensifs. En général, ils se traduisent par un changement dans l'équation d'états. Ce changement peut être soit paramétrique soit structurel/fonctionnel. Ces défauts induisent une instabilité de système.

Les défauts peuvent être classés selon leurs évolutions temporelles [Isermann *et al.*, 2002]:

- *Abrupt*: ce type de défaut se caractérise par un comportement temporel discontinu. Cette évolution, si elle ne correspond pas aux évolutions dynamiques normales attendues pour la variable (changement de consigne), est caractéristique d'une panne brutale de l'élément en question: arrêt total ou partiel de connexion. Une représentation mathématique de ce défaut est donné par:

$$f(t - t_{f_i}) = \begin{cases} \delta & t > t_{f_i} \\ 0 & t < t_{f_i} \end{cases}$$

Où  $f(t - t_{f_i})$  est le comportement temporel du défaut et  $\delta$  est un seuil constant.

- *Intermittent*: c'est un défaut caractéristique de faux contact ou de panne intermittente des capteurs. Ce type de défaut est un cas particulier de défaut abrupt avec la propriété particulière qu'il revient de façon aléatoire à sa valeur normale.
- *Graduel*: c'est un défaut caractéristique d'une usure d'une pièce ou d'un encrassement. Il est très difficile à être détecté à cause de son évolution temporelle qui peut être confondue à une modification paramétrique lente représentant la non stationnarité du procédé. Son évolution au cours du temps peut être exprimée par cette relation :

$$f(t - t_{f_i}) = \begin{cases} \delta \left(1 - e^{-\alpha(t-t_{f_i})}\right) & t \geq t_{f_i} \\ 0 & t < t_{f_i} \end{cases}$$

Où  $\alpha$  et  $\delta$  sont deux constantes positives.

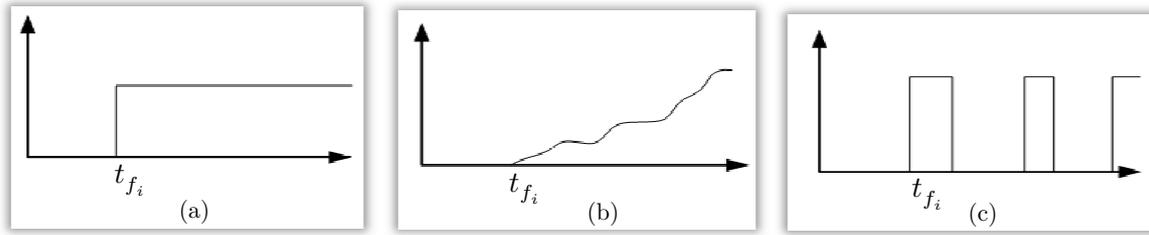


Figure 2. 4. Evolution temporelle d'un défaut: a) abrupt, b) graduel et c) intermittent.

Les défauts aussi peuvent être classés en défauts additifs et défauts multiplicatifs, selon leurs effets sur les performances du système (figure (2.5)). Les défauts additifs sont des signaux parasites qui viennent se superposer en un point du schéma fonctionnel. Les défauts capteurs et actionneurs sont généralement modélisés comme étant des défauts additifs, cependant, les défauts composants sont modélisés par des défauts multiplicatifs. Ces derniers induisent des changements sur la corrélation du signal de sortie du système, ainsi que des changements dans les caractéristiques spectrales et de la dynamique de système.



Figure 2.5. Classification des défauts : multiplicatifs et additifs.

### 2.2.2. Classification des méthodes de diagnostic

Les méthodes de diagnostic se distinguent selon différents critères: la dynamique de procédé (discret, continu ou hybride), sa complexité, l'implémentation de diagnostic en ligne et/ou hors ligne, la nature d'information (qualitative et/ou quantitative), sa profondeur (structurale, fonctionnelle et/ou temporelle), sa distribution (centralisée, décentralisée ou distribuée)...

Dans ce contexte, plusieurs classifications sont proposées dans la littérature [Franck et Köppenseliger, 1997 ; Isermann, 2006 ; Chantler et al., 1998 ; Dash et Venkatasubramanian, 2000 ; Venkatasubramanian et al., 2003]. Ces classifications sont influencées par les terminologies et les contextes particuliers de chaque communauté et ne sont pas toujours homogènes.

L'objectif est de déterminer la méthode la plus appropriée pour la résolution de notre problème de détection et diagnostic, ainsi que, positionner la classe des méthodes de diagnostic qui nous intéresse parmi les différentes méthodes de la littérature. Nous proposons une classification non exhaustive des méthodes de diagnostic en deux grandes familles (figure (2.6)):

- Les méthodes sans modèle mathématique qui ne nécessitent pas de connaissances accrues du système physique, mais utilisent des connaissances superficielles,
- Les méthodes à base de modèles qui nécessitent une connaissance approfondie du système physique.

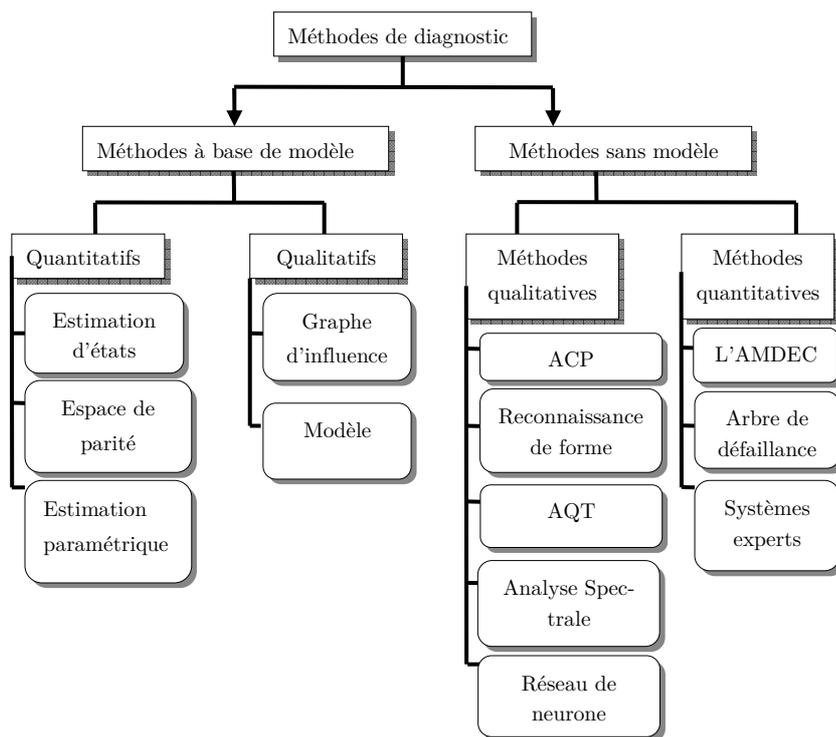


Figure 2.6. Une classification générale des méthodes de diagnostic.

### a. Méthodes sans modèle

Pour certaines applications industrielles, la conception d'un modèle mathématique est difficile, voire impossible à obtenir, à cause des nombreuses reconfigurations intervenant dans le processus de production ou de la complexité des phénomènes mises en jeu. Dans ce cas, on a recours à des méthodes qui ne nécessitent aucune connaissance approfondie du procédé. Deux classes, dans ce type d'approches, peuvent se présenter:

- Les méthodes quantitatives ou appelés méthodes à base de connaissances.
- Les méthodes qualitatives ou méthodes basées sur le traitement de données.

#### a.1. Méthodes qualitatives

Les méthodes qualitatives consistent en l'exploitation d'une base de connaissance symbolique et nécessitent l'existence d'un large éventail de données historiques correspondant aux divers modes de fonctionnement de l'installation.

- *Analyse en composantes principales*: l'analyse en composantes principales (ACP) est une technique statistique multi-variable, capable de compresser les données et de réduire leur taille. Elle peut être vue comme une technique de projection orthogonale linéaire qui projette les observations multidimensionnelles représentées dans un espace de dimension  $n$  ( $n$  est le nombre de variables observées) dans un sous espace de dimension  $q < n$ , en maximisant la variance des projections (ou en minimisant l'erreur quadratique d'estimation). Cette méthode est utilisée avec succès dans les études de diagnostic de [Jolliffe, 1986 ; Pessel et al., 2007 ; Oja et al., 1992].

- *Reconnaissance de forme*: l'objectif d'une méthode de reconnaissance de forme est la classification automatique d'objet suivant sa ressemblance par rapport à un objet de référence. Dans un problème de diagnostic, une classe est formée par l'ensemble d'observations caractérisant une situation ou un mode de fonctionnement de processus: par exemple, la classe  $C_1$  peut être liée au fonctionnement normal du procédé, la classe  $C_2$  pour le fonctionnement dégradé et la classe  $C_3$  pour le fonctionnement défaillant. Le diagnostic consiste à associer toute nouvelle observation à une classe. Le problème de diagnostic est équivalent à la recherche des frontières entre les classes qui minimisent l'erreur de classification. Le calcul de la distance (distance euclidienne) peut être choisi comme un critère de décision pour assigner une forme à une classe et de déterminer avec quelle confiance est effectuée cette décision.
- *Analyse qualitative des tendances*: son objectif est la représentation et l'analyse qualitative des signaux à travers l'information des dérivées de premier et de second ordre. Cette analyse se base sur la segmentation des signaux en unités appelées épisodes, en fonction de leur convexité et de leur pente. Elle est constituée de deux étapes de base: l'identification de tendances dans les mesures (ce processus doit être robuste par rapport au bruit en capturant seulement les variations importantes), et l'interprétation de ces tendances en termes de scénarios.
- *Analyse spectrale*: sous certaines conditions normales, certaines mesures ont un spectre typique de fréquence; toute déviation des caractéristiques fréquentielles d'un signal est reliée à une anomalie. Cette méthode s'avère intéressante pour l'analyse des signaux qui montrent des oscillations avec des périodes longues (les débits, les pressions...).  
L'application d'une procédure de décision permet de détecter et de localiser l'élément défaillant du système. Parmi les procédures de décision appliquées sur un échantillon de mesures on peut citer: le test empirique de franchissement de seuil, test de variance, le test de la moyenne ... [Basseville, 1988; 1997].
- *Réseaux de neurones*: le réseau de neurones peut être assimilé à un système informatique constitué d'un nombre de processeurs élémentaires interconnectés entre eux qui traite l'information qui lui arrive à partir des signaux extérieurs. Il est capable d'approcher une fonction non-linéaire dont on ne connaît que quelques points. De manière générale, l'utilisation d'un réseau de neurones se déroule en deux étapes. La première phase est l'apprentissage. Cet apprentissage permet alors, sur la base de l'optimisation d'un critère, de reproduire le comportement du système à modéliser. Il consiste dans la recherche d'un jeu de paramètres (poids) et peut s'effectuer de deux manières : supervisé (le réseau utilise les données d'entrée et de sortie du système à modéliser) et non supervisé (seules les données d'entrée du système sont fournies et l'apprentissage s'effectue par comparaison entre exemples). Une fois l'apprentissage est achevé, il peut être utilisé pour la généralisation. Il s'agit de la phase de test où de nouveaux exemples sont présentés au réseau de neurones pour juger de sa capacité à prédire les comportements du système modélisé. Lors d'un problème de diagnostic, leur emploi peut alors se faire à trois niveaux : 1) comme modèle du système à surveiller en état normal et générer un résidu d'erreur entre les observations et les prédictions, 2) comme système de détection en une seule étape (en tant que classificateur), ou 3) comme système d'évaluation de résidus pour le diagnostic.

### a.2. Méthodes quantitatives

Les méthodes quantitatives ou à base de connaissances sont mises en œuvre lorsque la majorité des mesures sont indisponibles et lorsque la construction du modèle s'avère difficile. Elles peuvent être utilisées pour identifier les causes des défaillances d'un procédé industriel. Il s'agit d'analyses fonctionnelles et structurelles qui se basent sur l'expérience et la connaissance de l'opérateur.

- *L'AMDE (Analyse des Modes de Défaillances et de leurs Effets) et ses dérivées*: c'est une méthode issue du contexte de la sûreté de fonctionnement. Villemeur dans [Villemeur, 1988] décrit l'AMDE comme une approche d'analyse consistant à identifier de façon inductive et systématique les modes de défaillances d'un système puis à en rechercher les causes et leurs effets. La démarche consiste à examiner comment et pourquoi les fonctions du système étudié risquent de ne plus être assurées correctement, en définissant les modes de défaillances potentielles, en recherchant les causes possibles de l'apparition de chaque mode de défaillance et enfin en étudiant pour chaque combinaison cause-mode de défaillance, les effets sur le système et sur l'utilisateur.

Une fois complétée par une analyse de criticité, l'AMDE devient l'AMDEC (Analyse des Modes de Défaillance et de leurs Effets et de la Criticité). Elle permet d'extraire les modes de défaillances les plus critiques. Elle consiste à examiner la criticité des modes de défaillances à partir de trois critères : gravité, fréquence d'apparition et probabilité de non-détection.

- *Arbre de causes à effets ou arbre de défaillances*: l'arbre de défaillances apparut la première fois dans [Villemeur, 1988], représente l'un des outils majeurs d'analyse de risques technologiques. L'objectif de cette approche est de déterminer les diverses combinaisons possibles d'évènements qui entraînent la réalisation d'un évènement indésirable unique. Sa représentation graphique est composée d'une structure arborescente permettant un traitement à la fois qualitatif et quantitatif. L'arbre de défaillances est composé de plusieurs niveaux où la racine correspond à l'évènement indésirable. Les niveaux sont ordonnés de manière successive tel que, chaque évènement est généré à partir des évènements du niveau inférieur par l'intermédiaire d'opérateurs logiques (ET, OU). La décomposition s'arrête aux niveaux des évènements élémentaires, caractérisés par le fait qu'ils sont indépendants et non décomposables en éléments plus simples.
- *les systèmes experts*: le système expert est un système informatique destiné à résoudre un problème précis à partir d'une analyse, d'une représentation des connaissances et du raisonnement d'un ou de plusieurs spécialistes de ce problème. Ils utilisent une information heuristique pour lier les symptômes aux défauts. A partir de l'ensemble des symptômes à sa disposition, il déduit toutes les conclusions possibles, élabore de nouvelles hypothèses et approfondit son diagnostic en exploitant des informations supplémentaires collectées sur le système à diagnostiquer. Les systèmes experts sont composés de deux parties indépendantes : une base de connaissances et un moteur d'inférence. La base de connaissances peut être subdivisée en deux groupes: la base de faits et la base de règles. La base de faits représente l'état du système observé. La base de règles est un ensemble de règles logiques qui permettent de déduire, d'inférer de nouveaux faits à partir d'autres faits déjà établis dé-

nommés les prémisses. Le moteur d'inférence a la lourde tâche de déduire tous les faits susceptibles d'être déduits à partir de la base de connaissances.

### *b. Méthodes à base de modèle*

L'utilisation des modèles pour le diagnostic date du début des années 70. Depuis de nombreux travaux ont été proposés [Willsky, 1976 ; Chow et Willsky, 1984 ; Basseville, 1988 ; Patton et al., 1989 ; Gentil, 1997 ; Isermann et Balle, 1997 ; Fussel et Isermann, 1998]. Une étude complète sur les méthodes à base de modèles peut être trouvée dans l'article [Frank, 1996 (b)] ou des livres récents comme [Patan, 2008 ; Chiang et al., 2001]. Ces méthodes sont des alternatives de la redondance physique ou matérielle. La structure générale de la plupart de ces méthodes se fonde sur l'idée de la redondance analytique [Chow et Willsky, 1984]. Le principe des méthodes à base de modèles est d'identifier l'écart entre le système réel et son modèle. Ces méthodes s'appuient sur des modèles comportementaux explicites du système soumis au diagnostic. Ces modèles peuvent être classés en deux catégories quantitatives ou qualitatives. Les méthodes à base de modèles sont plus performantes que celles basées sur les traitements de données ou sur les approches relationnelles [Frank, 1998]:

- le coût de développement et de maintenance est faible,
- les symptômes et les défauts ne doivent pas être anticipés,
- il s'agit des connaissances de conception plutôt que d'exploitation,
- la connaissance du système est découplée de la connaissance de diagnostic,
- seule l'information du comportement normal est prise en compte.

#### *b.1. méthodes quantitatives*

Ces méthodes reposent sur les relations mathématiques qui existent entre les variables. Ces modèles sont développés en utilisant les lois fondamentales de la physique (bilan de masse, d'énergie, de quantité de mouvement, ..) ou des relations de type entrée sortie. Selon Frank, les méthodes de diagnostic à base de modèles quantitatifs se décomposent en trois grands groupes [Frank, 1998]: les approches par espace de parité, les approches d'estimation paramétrique, et les approches d'estimation d'états.

- *Espace de parité* : l'objectif de cette méthode est la vérification de parité (la cohérence) des modèles du procédé avec les mesures issues des capteurs et des entrées connues (consignes, signal de commande,..). Les premières évocations et utilisations de l'approche par espace de parité ont été réalisées à partir des relations analytiques statiques [Evans, 1970 ; Potter et Suman, 1977 ; Daly et al., 1979]. Ces notions ont été ensuite généralisées par [Mironovski, 1980], puis par [Chow et Willsky, 1984]. Dans le cadre général, l'approche consiste à réaliser une redondance analytique (en mode temporel ou fréquentiel) entre les entrées et les sorties du système et cela indépendamment des états du système. La matrice de parité est définie comme étant une matrice de projection de la matrice d'observabilité, ce qui revient à éliminer l'influence des états sur les résidus.
- *Estimation paramétrique*: Le principe de l'estimation paramétrique est d'estimer en continu les paramètres du système, en utilisant les mesures d'entrées et de sorties. Le résidu est

alors la différence entre cette estimation et les valeurs de référence de l'état normal du procédé. Les premières applications de cette approche ont été faites par [Isermann, 1984].

Cette technique a été appliquée pour la détection des défauts pour les systèmes non-linéaires [Isermann, 2006]. De nombreuses applications ont été réalisées dans le domaine de l'automobile [Constantinescu et al., 1995 ; Dinca et al., 1999], ou pour un modèle non-linéaire d'un satellite [Jiang et al, 2008].

- *Estimation d'états*: Le principe de base de diagnostic par observateur consiste à estimer la sortie du système à partir des connaissances de la commande (ou partie de la commande), et des mesures (ou partie des mesures). Le vecteur résidu est ainsi construit comme l'écart entre la sortie estimée et la sortie mesurée. Ces résidus doivent servir d'indicateurs fiables du comportement du processus.

Plusieurs techniques d'observations ont été appliquées au diagnostic:

- *Filtre de détection de défaut*: cette méthode a été initiée par [Beard, 1971], puis a été formalisée par [Massoumnia, 1986] en 1986. L'objectif de cette approche est de construire des espaces de détection associés à chacun des défauts. Différents travaux traitent les méthodes des filtres de détection, parmi elles, retenons les travaux (Edelmayer et al., 1996 ; Frank et al., 1999 ; Rank et Niemann., 1999; Zhong et al., 2003...) pour le diagnostic robuste.
- *observateur de détection de défaut*: plusieurs travaux ont été menés dans le cadre des systèmes linéaires: les observateurs à entrées inconnues vise à annuler l'erreur d'estimation d'une ou plusieurs combinaisons linéaires des états du système étudié, et ce, indépendamment de la présence d'entrées inconnues [Chen et al., 1996]. Une première utilisation des observateurs à entrées inconnues pour le diagnostic remonte à [Viswanadham et al., 1987]. [Magni et Mouyon, 1994] ainsi que [Staroswiecki et al., 1991] apportent des précisions sur la mise en place d'un observateur pour la détection de défauts. Dans le cadre des systèmes non-linéaires: [Adjallah et al., 1993] ainsi que [Hammouri et al., 1999 (a) ; 1999 (b)] synthétisent des observateurs pour des systèmes non-linéaires. La détection de défaut à base d'observateur non-linéaire constitue l'objectif principal de cette thèse. Une brève description des méthodes de génération des résidus par observateur pour des systèmes non-linéaires fera l'objet de la section (2.2.3).

## b.2. Méthodes qualitatives

La communauté d'Intelligence Artificielle (IA) a proposé un raisonnement qualitatif (ou semi-qualitatif) basé sur l'établissement de relations causes à effets. En effet, le diagnostic est typiquement un système causal puisqu'il consiste à établir des hypothèses sur les composants défaillants qui sont l'origine du dysfonctionnement observé. Le raisonnement qualitatif exprime le lien entre un composant et les formules décrivant son comportement. Parmi les méthodes les plus utilisées, on peut citer:

- *Les graphes causaux (ou graphes d'influence)*: ils déterminent les composants défectueux qui peuvent expliquer le fonctionnement anormal observé. Le diagnostic basé sur un graphe causal consiste à rechercher la variable source dont la déviation est suffisante pour expli-

quer toutes les déviations détectées sur d'autres variables [Travé-Massuyès et al., 97]. Deux principaux types de structure causale sont proposés: le premier type lie la causalité aux équations décrivant le système (analyse globale) [Pons et Travé-Massuyès, 1997; Blanke et al., 2006] tandis que le deuxième axe lie la causalité à la structure du système (analyse locale).

- *La logique floue*: la logique floue est une théorie mathématique introduite par [Zadeh, 1965], qui permet de tenir compte des incertitudes et permet une fusion des informations. L'idée de l'approche floue est de construire un dispositif, appelé système d'inférences floues, capable d'imiter les prises de décision d'un opérateur humain à partir des règles verbales traduisant ses connaissances relatives à un processus donné. Trouver une relation mathématique entre un défaut et ses symptômes s'avère souvent difficile. Cependant, en s'appuyant sur leur expérience, les opérateurs humains sont capables de déterminer l'élément défaillant qui est à l'origine des symptômes observés. Ce type de savoir peut être exprimé à l'aide de règles de la forme : SI *condition* ALORS *conclusion*. Où la partie *condition* comporte les symptômes observés et la partie *conclusion* l'élément défaillant. Ainsi le problème de diagnostic est considéré comme un problème de classification. Le vecteur des symptômes du classificateur, élaboré à partir des grandeurs mesurées sur le système, peut être vu comme une forme, qu'il s'agit de classer parmi l'ensemble des formes correspondant à un fonctionnement normal ou non.

### *c. Analyses comparatives des méthodes de diagnostic*

Comparer les méthodes de diagnostic est une tâche très difficile. En effet, la décision de choix d'une méthode de détection de défaut dépend de plusieurs facteurs: connaissance a priori disponible sur le système, présence ou non d'un modèle mathématique de processus à diagnostiquer, type de défauts à détecter, présence ou non d'entrées inconnues (bruit, incertitude), non-linéarité, système en boucle fermé ou ouvert...

Par exemple, pour les systèmes électriques et mécaniques, il est simple de concevoir un modèle mathématique. Le choix des méthodes de diagnostic à base de modèles sera préféré. Par contre, pour les processus industriels chimiques, la modélisation est difficile à mener car même si un modèle mathématique est obtenu, il est complexe. Dans ce cas, les méthodes de détection de défauts sans modèle de référence peuvent être appliquées. [Isermann, 2006] a proposé une étude comparative entre les méthodes de diagnostic par redondance analytique. Il est à signaler que de nombreuses méthodes sont nées du couplage de plusieurs stratégies de diagnostic dont le but est de combiner les avantages de chaque stratégie. Quelques approches combinées sont présentées dans [Isermann, 2006].

Dans notre cas, l'intérêt des méthodes utilisant des modèles mathématiques (section 2.2.2 (a)) est plus grand que celui n'en utilisant pas (section 2.2.2 (b)), et cela du fait de l'existence d'un modèle des stations d'épuration biologiques par boues activées (équation (3.9)). Parmi les méthodes de diagnostic à base de modèle mathématique: approche par estimation d'état et approche par estimation paramétrique. Ces deux approches se complètent et s'opposent à la fois. D'après ce qui précède, les méthodes d'estimation paramétrique donnent confiance aux données provenant des capteurs et d'actionneurs et met en défaut les paramètres du modèle. En revanche, pour les méthodes par estimation d'état les paramètres du modèle sont supposés corrects et les défauts sont dus aux capteurs et actionneurs. Ainsi le choix entre ces deux méthodes

est difficile. Nous pouvons nous diriger vers une approche d'estimation d'état en supposant que les paramètres soient connus et que les défauts soient expliqués par un dérèglement des mesures des capteurs et actionneurs. Dans le cas contraire, nous nous dirigeons vers une approche par estimation paramétrique.

Les méthodes à base d'observateur et celles par espace de parité diffèrent dans la manière de générer des résidus. L'approche espace de parité est une méthode de génération de résidus directe, basée sur une élimination des inconnues à l'aide de la redondance. Par contre, la génération est indirecte à l'aide de méthodes à base d'observateur basée sur le calcul de l'erreur d'estimation. [Frank, 1990] a prouvé sous certaines hypothèses l'équivalence entre l'espace de parité et les résidus utilisant des observateurs. L'approche par espace de parité, utilise une matrice, appelé matrice de parité. La difficulté principale de l'approche par espace de parité réside dans la recherche de la matrice de parité.

Après l'exposition des méthodes de diagnostic, nous retiendrons les méthodes de diagnostic à base de modèles quantitatifs et en particulier nous nous concentrons sur les méthodes à base d'observateurs.

Le tableau (2.1) propose une analyse comparative synthétique des approches de détection et de diagnostic que nous venons d'exposer.

### **2.2.3. Détection, localisation et identification des défauts à base d'observateur non-linéaire**

Les systèmes ont en général un comportement non-linéaire. Une solution pour remédier à la complexité des outils de FDI aux systèmes non-linéaires est de se ramener au cas linéaire (en utilisant par exemple un modèle linéaire autour d'un point de fonctionnement) afin d'utiliser les méthodes de FDI linéaire qui sont bien maîtrisées. Néanmoins, ce type de méthodes soulève des problèmes de généralité. En effet, considérons le cas où le générateur de résidus se base sur un modèle linéarisé autour d'un point de fonctionnement. Lorsque l'état s'écarte sensiblement de ce point de fonctionnement, des dérives importantes peuvent être observées, à cause de ce caractère non-linéaire du comportement [Maquin et al., 1997]. Le principal inconvénient de ces méthodes est qu'elles ne peuvent pas s'appliquer que sous des conditions très restrictives [Xu, 2002]. De plus elles peuvent augmenter les taux de fausses alarmes [Frank et Ding, 1997]. Ces limitations des techniques linéaires de diagnostic ont motivé les recherches des techniques de FDI non-linéaires représentant au mieux le comportement du processus. Parmi ces approches nous retiendrons celles qui se basent sur les modèles mathématiques puisqu'elles offrent un grand nombre d'informations sur le système. Comme indiqué dans le paragraphe 2.2.2 (b), parmi ces méthodes, on a l'estimation d'état, l'espace de parité et l'estimation des paramètres. Un intérêt particulier sera accordé aux méthodes à base d'observateurs.

Dans la section suivante, une étude bibliographique sur les méthodes de détection et de générations des résidus à base d'observateurs non-linéaires est donnée. Ainsi une présentation des méthodes d'évaluation des résidus est évoquée par la suite.

Cette méthode de diagnostic est composée de trois fonctions de base: la génération des résidus, la détection, la localisation et éventuellement l'identification.

Modèle	Type	Méthodes	Avantages	Inconvénients
Sans modèle	Méthodes quantitatives	Les arbres de défaillances : l'AMDE et l'AMDEC	<ul style="list-style-type: none"> <li>- Démarche très puissante pour l'analyse de défaillances simples d'éléments conduisant à la défaillance globale</li> <li>- Aucun traitement nécessaire</li> <li>- Très efficace pour l'analyse et la résolution des problèmes de diagnostic de défaillances des procédés industriels.</li> <li>- Identification des modes communs de défaillances pouvant affecter le système</li> </ul>	<ul style="list-style-type: none"> <li>- Demande une longue expérience</li> <li>- Toute modification ou évolution du système nécessite la réécriture du tableau</li> <li>- Mal adaptée aux systèmes dynamiques, dépendant fortement du temps, à cause du nombre élevé de variables, processus,</li> <li>- Incapacité de traitement des défauts multiples.</li> <li>- Arbre construit dépendant de son créateur : sensibilité aux erreurs</li> <li>- Pas de méthode formelle pour vérifier l'exactitude de l'arbre développé.</li> </ul>
		Les systèmes experts	<ul style="list-style-type: none"> <li>- Temps de calcul faible</li> <li>- aucun raisonnement compliqué (pas de calcul intermédiaire) : génération directe des règles par un expert humain qui sont des résultats compréhensible par l'opérateur.</li> </ul>	<ul style="list-style-type: none"> <li>- Dépend de l'expertise faite sur le système</li> <li>- Difficultés d'acquisition de l'expertise</li> <li>- Manque de généricité car les règles dépendent de l'architecture du système</li> <li>- Problème de l'évolution du système : ajout/ changement de composant implique nouvelle expertise</li> <li>- Robustesse : règles fixées et non robustes face aux situations non reconnues</li> <li>- Données incertaines : difficultés dans l'analyse d'un ensemble de données non corrélées incomplètes et ambiguës</li> <li>- Manque de connaissances profondes : pas d'explications sur les conclusions adoptées.</li> <li>- Incohérence des règles : ajout ou suppression de règle → impact sur autres règles difficile à détecter</li> </ul>
	Méthodes qualitatives	L'analyse en composantes principales	<ul style="list-style-type: none"> <li>- Outil puissant capable de réduire la taille des données de telle sorte que l'information soit conservée</li> <li>- Facilite l'analyse des données</li> <li>- Manipulation du bruit et corrélation pour extraire l'information efficacement</li> </ul>	<ul style="list-style-type: none"> <li>- Représentation invariante dans le temps</li> <li>- Mise à jour périodique</li> <li>- Aucune propriété de signature ce qui rend difficile l'isolation de fautes</li> </ul>

Sans modèle	Méthodes qualitatives	Reconnaissance de forme	<ul style="list-style-type: none"> <li>- Simple à appliquer</li> <li>- Temps de calcul faible pour la classification d'une nouvelle observation et indépendant de la dimension de l'ensemble d'apprentissage</li> <li>- Capable de traiter des données qui sont à la fois incertaines et imprécises.</li> </ul>	<ul style="list-style-type: none"> <li>- Pas de règles générales pour le choix de l'espace de représentation qui est le facteur de réussite de diagnostic</li> <li>- Nombre de classes supposé connu au départ → connaissance supposée exhaustive</li> <li>- Toute observation doit appartenir à une classe définie</li> <li>- Les caractéristiques de certains modes de fonctionnement restent inconnues</li> </ul>
		L'analyse fréquentielle	<ul style="list-style-type: none"> <li>- Utile pour analyser des signaux qui montrent des oscillations avec des périodes longues (débit courant électrique, pression, ...)</li> <li>- Relativement simple à mettre en place</li> </ul>	<ul style="list-style-type: none"> <li>- Sensibilité aux bruits de mesure quand ceux-ci coïncident avec la zone fréquentielle d'intérêt</li> <li>- Nécessité d'un échantillonnage fréquent pour la reconstitution du signal tout en minimisant les pertes de fréquence</li> </ul>
		Les réseaux de neurones artificiels	<ul style="list-style-type: none"> <li>- Faible sensibilité aux bruits de mesure,</li> <li>- Capacité à résoudre des problèmes non linéaires et multi-variables</li> <li>- Capacité à apprendre en ligne et en temps réel</li> </ul>	<ul style="list-style-type: none"> <li>- Temps de calcul important pendant la phase d'apprentissage</li> <li>- Définition des paramètres du réseau délicate</li> </ul>
		L'analyse qualitative de la tendance	<ul style="list-style-type: none"> <li>- Méthode adaptée aux incertitudes, incomplétude et hétérogénéité des données</li> <li>- Adaptée aux procédés complexes (grand volume de données) grâce à sa représentation symbolique</li> <li>- Représentation de la connaissance adaptée à l'interprétation de l'opérateur humain</li> </ul>	<ul style="list-style-type: none"> <li>- Observation délicate des dérivées</li> <li>- Nécessite la mise en œuvre d'une étape de filtrage de bruit, d'un traitement de discontinuités</li> <li>- Nécessite une description exhaustive des modes du procédé et des comportements de défauts</li> <li>- Approche inappropriée dans le cas de dérivées lentes</li> </ul>
	Modèle Quantitatif	L'espace de parité	<ul style="list-style-type: none"> <li>- Méthode très générale</li> <li>- Facilite l'analyse des données : isolation de la défaillance</li> <li>- Connaissance du système découplée de la connaissance du diagnostic</li> <li>- Moindre coût de développement</li> <li>- Structure intéressante pour les procédés complexes</li> </ul>	<ul style="list-style-type: none"> <li>- Prise en compte d'incertitudes additives</li> <li>- Mal adaptée pour les non linéarités et les pannes non additives</li> </ul>

<b>Avec modèle</b>	Modèle Quantitatif	L'estimation paramétrique	<ul style="list-style-type: none"> <li>- Apporte de l'information sur l'ampleur des déviations</li> <li>- Méthode bien adaptée aux défauts multiplicatifs (ayant une incidence sur les paramètres)</li> <li>- Application à des procédés à faible nombre de variables où des modèles précis peuvent être définis.</li> </ul>	<ul style="list-style-type: none"> <li>- Nécessité d'avoir un système physique excité en permanence : Problème dans le cas de procédés dangereux ou fonctionnant en mode stationnaire</li> <li>- Relations entre les paramètres mathématiques et physiques pas toujours inversibles de façon unitaire → tâche de localisation du diagnostic difficile</li> <li>- Utilisation difficile pour des installations complexes due au grand nombre de variables mis en jeu</li> </ul>
		L'estimation d'état	<ul style="list-style-type: none"> <li>- Applicable aux systèmes linaires et non linéaires</li> <li>- Méthodes très répandues</li> <li>- Calcul simple</li> <li>- Possibilité de découpler les réponses des entrées inconnues</li> <li>- Méthode plus robuste aux bruits de mesures</li> </ul>	<ul style="list-style-type: none"> <li>- Nécessité d'avoir un modèle précis et complet</li> <li>- Mal adaptée au processus complexe</li> <li>- Erreur de diagnostic due aux perturbations</li> <li>- Pas de garantie de détection si le type de défaut n'a pas été modélisé</li> <li>- Adaptabilité difficile aux changements de processus et manque de méthode générale dû au caractère local du modèle (appliqué au système étudié)</li> </ul>
	Modèle qualitatif	Les graphes d'influence	<ul style="list-style-type: none"> <li>- Connaissance de comportements (modèles) de défaut non nécessaire pour la localisation</li> <li>- Capacité explicative</li> <li>- Exploitation au maximum de la structure</li> <li>- Langage graphique unique applicable à de nombreux systèmes</li> <li>- Repose uniquement sur le parcours de chemins causaux</li> </ul>	<ul style="list-style-type: none"> <li>- Génération d'un grand nombre d'hypothèses pouvant conduire à un diagnostic erroné</li> <li>- Construction de la structure peut entraîner des pertes d'information</li> </ul>
		La logique floue	<ul style="list-style-type: none"> <li>- Description de la structure physique connue d'un système</li> <li>- Approximation du comportement d'un système complexe</li> <li>- Variable de sortie directement liée aux variables d'entrée</li> </ul>	<ul style="list-style-type: none"> <li>- Règle n'ayant pas de signification sémantique claire</li> <li>- Ne fournit pas une meilleure compréhension de la relation entre les variables</li> </ul>

Tableau 2.1. Une analyse comparative synthétique des approches de détection et de diagnostic.

### a. Génération des résidus

La première étape d'une procédure de FDI est la génération des résidus. Ce terme de résidu est bien choisi et reflète d'une manière claire son utilité. En effet, c'est un indicateur de la défaillance du système. Dans le cas des observateurs, les résidus représentent l'écart entre la sortie réelle et la sortie estimée. En d'autres termes, c'est l'erreur d'estimation.

Les approches par observateurs sont les méthodes les plus couramment utilisées pour la génération des résidus [Chen et Patton, 1999 ; Frank et Ding, 1997 ; Witczak, 2007; Garcia et Frank, 1997 ; Frank, 1994; 1996 (a)]. Quelques unes sont décrites par la suite:

#### Observateur de Luenberger étendu

[Clark et al., 1975] ont utilisé l'observateur de Luenberger pour la détection de défaut des systèmes linéaires. Pour la détection des défauts des systèmes non-linéaires, on peut linéariser le modèle non-linéaire autour d'un point de fonctionnement et appliquer par la suite l'observateur de Luenberger. Le modèle non-linéaire est donc approché autour d'un point de fonctionnement par un modèle linéaire tel que sa dynamique soit proche de celle du modèle non-linéaire dans un voisinage du point de linéarisation considéré. Cependant, si le système s'éloigne de ce point de fonctionnement, le modèle linéaire présente une dynamique moins proche du modèle non-linéaire et l'erreur faite par la linéarisation est plus grande. L'idée d'un observateur de Luenberger étendu est de linéariser le modèle autour de la trajectoire estimée  $\hat{x}(t)$ , au lieu de la linéarisation autour d'un point de fonctionnement (exemple  $x = 0$ ), puis d'utiliser un observateur de Luenberger par la suite.

Considérons le système non-linéaire suivant:

$$\begin{aligned}\dot{x} &= a(x, u), & x(0) &= x_0 \\ y &= c(x, u)\end{aligned}\tag{2.1}$$

Les équations de l'évolution d'un observateur de Luenberger étendu s'écrivent :

$$\begin{aligned}\dot{\hat{x}} &= a(\hat{x}, u) + L(\hat{x}, u)(y - c(\hat{x}, u)), & \hat{x}(0) &= \hat{x}_0 \\ \hat{y} &= c(\hat{x}, u)\end{aligned}\tag{2.2}$$

Avec  $L(\hat{x}, u)$  le gain d'observateur qui est calculé de façon à ce que les valeurs propres de la quantité  $\left(\frac{\partial a(x, u)}{\partial x} - L(\hat{x}, u) \frac{\partial c(x, u)}{\partial x}\right)$  soient stable.

Une étude détaillée peut être trouvée dans [Zeitz, 1987]. Une approche similaire pour l'estimation de l'état et son application à la détection des défauts a été proposée dans [Adjallah, 1994]. À cause du calcul répétitif de gain à chaque instant, l'observateur de Luenberger étendu est rarement utilisé dans la pratique. Dans le cas stochastique, le filtre de Kalman étendu constitue l'approche la plus connue et la plus populaire pour traiter les systèmes non-linéaires. Il a été implémenté pour la première fois par Schmidt [Schmidt, 1966]. Ce filtre est similaire à l'observateur de Luenberger étendu. En effet, la linéarisation s'effectue au tour de la trajectoire estimée par le filtre (voir section 2.3.6).

### L'observateur de Thau

L'observateur de Thau a été développé par [Thau, 1973] pour l'estimation d'état d'une classe des systèmes non-linéaires. Par la suite, une application de cet observateur pour la détection de défauts a été proposée par [Schreier et al., 1997]. Les systèmes non-linéaires considérés sont de la forme:

$$\begin{aligned}\dot{x} &= Ax + Bu + g(x, u) \\ y &= Cx\end{aligned}\tag{2.3}$$

Où le couple  $(C, A)$  est supposé observable, et  $g(x, u)$  est une fonction non-linéaire vectorielle dépendante de  $x$  et  $u$  et satisfaisant la condition de Lipchitz :

$$\|g(x, u) - g(\hat{x}, u)\| \leq \gamma \|x - \hat{x}\|$$

La structure de l'observateur de Thau est la suivante :

$$\begin{aligned}\dot{\hat{x}} &= A\hat{x} + Bu + g(\hat{x}, u) + L(y - \hat{y}) \\ r &= y - C\hat{x}\end{aligned}\tag{2.4}$$

Avec le gain d'observateur  $L = P^{-1}C'$ , et  $P$  est la solution de l'équation de Lyapunov :

$$A'P + PA - C'C + \theta P = 0\tag{2.5}$$

Où  $\theta$  est une constante positive choisie de telle sorte qu'elle vérifie l'équation de Lyapunov précédente (2.5).

### Observateur non linéaire à entrées inconnues

Les observateurs à entrées inconnues annulent complètement l'effet des perturbations sur les résidus. L'observateur à entrées inconnues a d'abord été appliqué pour la détection de défauts dans [Wünnenberg et Frank, 1987] et plus tard des contributions ont été introduites dans [Frank et Wünnenberg, 1989] pour les systèmes linéaires. La démarche a été étendue à une classe de systèmes non linéaires dans [Wünnenberg, 1990]. Cette approche peut être appliquée à une classe de systèmes non linéaires modélisée par :

$$\begin{aligned}\dot{x} &= Ax + B(y, u) + E_d d + E_f(x, u) f_f \\ y &= Cx + F_f(x, u) f_s\end{aligned}\tag{2.6}$$

Où  $d$  représente les entrées inconnues,  $f_f$  et  $f_s$  représentent respectivement les défauts actionneurs et les défauts capteurs. Ainsi l'observateur est donné par:

$$\begin{aligned}\dot{\hat{x}} &= F\hat{x} + J(y, u) + Gy \\ r &= L_1\hat{x} + L_2y\end{aligned}\tag{2.7}$$

Pour assurer le découplage des entrées inconnues, les conditions suivantes doivent être satisfaites [Frank et al., 1999].

$$TA - FT = GC, \quad F \text{ stable}$$

$$J(y, u) = TB(y, u)$$

$$L_1T + L_2C = 0, TE_d = 0$$

$$\text{rang} \left( TE_f(x, u) \right) = \text{rang} \left( E_f(x, u) \right)$$

$$\text{rang} \left( \begin{bmatrix} G \\ L_2 \end{bmatrix} \right) = \text{rang} \left( F_f(x, u) \right)$$

Une fois les conditions ci-dessus remplies, on définit l'erreur d'estimation par  $e = Tx - \hat{x}$ , ainsi que la dynamique de l'erreur et le résidu est régis par:

$$\begin{aligned} \dot{e} &= Fe - GE_f(x, u)f_f + TF_f(x, u)f_s \\ r &= L_1e + L_2F_f(x, u)f_s \end{aligned} \quad (2.8)$$

Un inconvénient majeur des méthodes à base d'observateur à entrées inconnues est qu'elles ne sont applicables que pour une classe limitée de systèmes non-linéaires. D'autres méthodes ont été proposées dans la littérature pour transformer les modèles non-linéaires à la forme convenable pour l'application des observateurs à entrées inconnues, mais les conditions de transformation sont trop restrictives [Chen et Patton, 1999].

### Observateurs non-linéaires à découplage des perturbations

Pour remédier aux conditions très restrictives de l'existence des observateurs à entrées inconnues, [Seliger et Frank, 1991] ont proposé des observateurs non-linéaires à découplage des perturbations. Une classe plus large des systèmes non-linéaires est considérée pour cette approche.

$$\begin{aligned} \dot{x} &= a(x, u) + E_d(x)d + E_f(x, u)f \\ y &= c(x, u) \end{aligned} \quad (2.9)$$

Le découplage d'une partie de l'état du système de l'effet des perturbations, à l'aide d'une transformation non-linéaire  $z = T(x)$ , permet de mettre le système (2.9) sous la forme :

$$\dot{T}(x) = \frac{\partial T(x)}{\partial x} (a(x, u) + E_d(x)d + E_f(x, u)f) \quad (2.10)$$

La transformation est choisie de façon à ce que la sensibilité du système transformé aux perturbations soit minimale et maximale aux défauts. Une telle transformation peut être donnée par :

$$\frac{\partial T(x)}{\partial x} E_d(x) = 0 \quad (2.11)$$

Ainsi le système transformé n'est plus affecté par les perturbations et se met sous la forme :

$$\dot{z} = \frac{\partial T(x)}{\partial x} (a(x, u) + E_f(x, u)f_f) \quad (2.12)$$

$$y^* = c^*(z, u, y)$$

[Seliger et Frank, 1991] ont discuté le cas où la matrice de distribution des perturbations est dépendante de  $u$ . Dans ce cas la transformation requise sera dépendante en  $u$ .

### Les observateurs grand-gain

Les observateurs grand-gain sont des observateurs qui présentent d'excellentes propriétés globales dont la référence historique est l'article de [Gauthier et al, 1992]. Ils ont été appliqués pour la détection de défauts par [Hammouri et al., 1999 (a)]. Pour plus de détail sur ce type d'observateur voir section (2.3.6).

### Observateur à mode glissant

Les observateurs à mode glissant ont été largement utilisés pour la détection de défauts des systèmes linéaires [Edwards et al., 2000 ; Tan et Edwards, 2002] ainsi que pour les systèmes non-linéaires [Yan et Edwards, 2007 ; 2008]. La robustesse aux bruits et aux incertitudes les rend plus appropriés pour l'estimation d'états et le diagnostic. Le principe des observateurs à mode glissant consiste à contraindre, à l'aide d'une fonction discontinue, les dynamiques d'un système d'ordre  $n$  à converger vers une variété  $S$  de dimension  $(n - p)$  dite surface de glissement ( $p$  est la dimension du vecteur de mesure). L'attractivité de cette surface est assurée par des conditions appelées conditions de glissement. Si ces conditions sont vérifiées, le système converge vers la surface de glissement et y évolue selon une dynamique d'ordre  $(n - p)$ . Dans le cas des observateurs à mode glissant, les dynamiques concernées sont celles des erreurs d'observations  $e_1(t) = z_1(t) - \hat{z}_1(t)$ . A partir des valeurs initiales  $e_1(0)$ , ces erreurs convergent vers les valeurs d'équilibre en deux étapes :

Dans la première étape, la trajectoire des erreurs d'observation évolue vers la surface de glissement sur laquelle les erreurs de sortie  $e_y = y - \hat{y}$  sont nulles. Cette étape est appelée mode d'atteinte. L'attractivité de la surface de glissement est garantie si la fonction de Lyapunov  $V(x) = S'S$  vérifie la condition  $\dot{V}(x) < 0$  si  $S \neq 0$ .

Dans la deuxième phase, la trajectoire des erreurs d'observation glisse sur la surface de glissement avec des dynamiques imposées de manière à annuler toutes les erreurs d'observation. Ce dernier mode est appelé mode de glissement. Durant cette étape, le gain de correction agit de manière à satisfaire la condition d'invariance suivante :  $\{\dot{S} = 0 \text{ et } S = 0\}$ .

Nous décrivons par la suite les principales étapes de conception d'un observateur à mode glissant donné par [Yan et Edwards, 2007]. Considérons une classe des systèmes non-linéaires décrite par :

$$\begin{aligned} \dot{x} &= Ax + g(x, u) + E\psi(t, u, x) + Dd(t, u, y) \\ y &= Cx \end{aligned} \quad (2.13)$$

$x \in \mathbb{R}^n$ ,  $u \in \mathbb{R}^m$ , et  $y \in \mathbb{R}^p$  sont respectivement le vecteur d'état, d'entrées et de sorties.  $A \in \mathbb{R}^{n \times n}$ ,  $E \in \mathbb{R}^{n \times r}$ ,  $D \in \mathbb{R}^{n \times q}$  et  $C \in \mathbb{R}^{p \times n}$  sont des matrices constantes. La fonction inconnue  $g(x, u)$  est Lipchitzienne par rapport à  $x$ . Une fonction de perturbation inconnue  $\psi(t, u, x)$  est bornée et vérifie :

$$\|\psi(t, u, x)\| \leq \xi(t, u, x)$$

Si la condition  $\text{rang}(C[E \ D]) = \text{rang}([E \ D])$  est satisfaite, alors il existe un changement de coordonnées, tel que le triplet  $(A, [E \ D], C)$  a la structure suivante:

$$\left( \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}, \begin{bmatrix} 0_{(n-p) \times r} & 0_{(n-p) \times q} \\ E_2 & D_2 \end{bmatrix}, [0_{p \times (n-p)} \ C_2] \right) \quad (2.14)$$

Où  $A_1 \in \mathbb{R}^{(n-p) \times (n-p)}$ ,  $C_2 \in \mathbb{R}^{p \times p}$  sont des matrices non singulières. Ainsi le système (2.13) peut donc être mis sous la forme:

$$\begin{aligned} \dot{x}_1 &= A_1 x_1 + A_2 x_2 + g_1(x, u) \\ \dot{x}_2 &= A_3 x_1 + A_4 x_2 + g_2(x, u) + E_2 \psi(t, u, x) + D_2 d(t, u, y) \\ y &= C_2 x_2 \end{aligned} \quad (2.15)$$

Les zéros invariants de  $(A, [E \ D], C)$  sont dans  $\mathbb{C}^-$ . Il existe une matrice  $L$  de la forme  $L = [L_1 \ 0]$  tel que  $A_1 + LA_3$  est stable.

La reconstruction des variables d'états est basée sur les sorties mesurées. En appliquant le changement de coordonnées suivant, les sorties apparaissent directement comme des composantes du vecteur d'état.

$$z = \begin{bmatrix} I_{n-p} & L \\ 0 & I_p \end{bmatrix} x \quad (2.16)$$

On obtient

$$\begin{aligned} \dot{z}_1 &= (A_1 + LA_3)z_1 + (A_2 + LA_4 - (A_1 + LA_3)L)z_2 + [I_{n-p} \ L]g_1(T^{-1}z, u) \\ \dot{z}_2 &= A_3z_1 + (A_4 - A_3L)z_2 + g_2(T^{-1}z, u) + E_2\psi(T^{-1}z, u, t) + D_2d(t, u, y) \\ y &= C_2z_2 \end{aligned} \quad (2.17)$$

Ainsi, un observateur à mode glissant pour le système (2.13) est donné par:

$$\begin{aligned} \dot{\hat{z}}_1 &= (A_1 + LA_3)\hat{z}_1 + (A_2 + LA_4 - (A_1 + LA_3)L)C_2^{-1}y + [I_{n-p} \ L]g_1(T^{-1}\hat{z}, u) \\ \dot{\hat{z}}_2 &= A_3\hat{z}_1 + (A_4 - A_3L)\hat{z}_2 - K(y - C_2\hat{z}_2) + g_2(T^{-1}\hat{z}, u) + \vartheta(t, u, y, \hat{y}, \hat{z}) \\ \hat{y} &= C_2\hat{z}_2 \end{aligned} \quad (2.18)$$

La matrice de gain  $K$  est choisie de manière que  $C_2(A_2 - A_3L)C_2^{-1} + C_2K$  soit symétrique définie négative. La fonction  $\vartheta$  est de la forme :

$$\vartheta = k(\cdot)C_2^{-1} \frac{y - \hat{y}}{\|y - \hat{y}\|} \quad \text{si } y - \hat{y} \neq 0$$

Où  $k(\cdot)$  est scalaire positive à déterminer. Les erreurs d'estimation d'état et de sortie sont données par  $e_1 = z_1 - \hat{z}_1$  et  $e_y = y - \hat{y} = C_2(z_2 - \hat{z}_2)$ . La dynamique des erreurs d'estimation s'écrit:

$$\begin{aligned} \dot{e}_1 &= (A_1 + LA_3)e_1 + [I_{n-p} \ L](G(T^{-1}z, u) - G(T^{-1}\hat{z}, u)) \\ \dot{e}_y &= C_2A_3e_1 + (C_2(A_4 - A_3L)C_2^{-1} + C_2K)e_y + C_2(G_2(T^{-1}z, u) - G_2(T^{-1}\hat{z}, u)) - C_2\vartheta \\ &\quad + C_2D_2d(y, u, t) + C_2E_2\psi(T^{-1}z, u, t) \end{aligned}$$

Considérons la surface de glissement  $S = \{(e_1, e_y) | e_y = 0\}$

$e_1$  est asymptotiquement stable si l'inégalité matricielle  $\bar{A}'\bar{P}' + \bar{P}\bar{A} + \frac{1}{\epsilon}\bar{P}\bar{P}' + \epsilon\gamma_g^2I_{n-p} + \alpha P < 0$  se résout pour  $\bar{P}$ . Où  $\bar{P} = P[I_{n-p} \ L]$ ,  $\bar{A} = [A_1 \ A_3]^T$ ,  $P > 0$ ,  $\alpha$  et  $\epsilon$  sont des constantes positives et  $\gamma_g$  est une constante Lipchitz de  $g(x, u)$  par rapport à  $x$ .

A ce stade, le mode de glissement associé à la surface de glissement  $S$  est stable. L'objectif de ce qui suit est de déterminer le gain  $k(\cdot)$  de façon à ce que la surface de glissement soit atteinte au bout d'un temps fini  $t_s$  et le mode de glissement soit maintenu. Pour cela, on considère l'équation candidate de Lyapunov suivante :  $\tilde{V}(e_y) = e_y'e_y$ . En choisissant  $k(\cdot)$  qui vérifie  $k(t, u, y, \hat{z}) \geq (\|C_2A_3\| + \|C_2\|\gamma_g + \|C_2E\|\gamma_\xi)\hat{w} + \|C_2E_2\|\xi(T^{-1}\hat{z}, u, t) + \|C_2D_2\|\rho(y, u, t) + \eta$ , la dérivée de l'équation de Lyapunov précédente satisfait  $\dot{\tilde{V}} \leq -2\eta\sqrt{\tilde{V}}$ . Où  $\eta$  est une constante positive et  $\hat{w}$  est la solution de l'équation différentielle  $\dot{\hat{w}} = -\frac{1}{2}\alpha\hat{w}(t)$ . Par conséquent, le régime glissant est établi dans un temps fini  $t_s$ .

### b. La détection

Après la génération des résidus, la deuxième étape du système FDI est la détection. Elle permet de déterminer la présence ou non d'un défaut. La détection consiste en l'évaluation des résidus. D'une manière générale, les résidus sont directement liés aux variations provenant des sorties (bruit de mesures, défauts capteurs), des entrées (défaut actionneurs, entrées inconnues, perturbations...)

Soit un vecteur de mesure  $y$  (de dimension  $p$ ) lié au vecteur d'état  $x$  (de dimension  $n$ ), au vecteur de commande  $u$  (de dimension  $m$ ), au vecteur de paramètre  $\theta$  (de dimension  $q$ ), au vecteur de bruit  $\vartheta$  (de dimension  $b$ ) et au vecteur des défauts  $d$  (de dimension  $l$ ) par une fonction relative au comportement de processus, on peut définir :

$$y = f(x, u, \theta, \vartheta, d)$$

Un résidu est un vecteur dont la dimension dépend à la fois de la méthode utilisée par la génération et du système étudié. Un résidu dans le cas d'estimation de sortie est donné par :

$$r = y - f(x, u, \theta, \vartheta, d)$$

Avec  $r = 0$  dans le cas sans défaut  $d$  et sans bruit  $\vartheta$ .

Comme il est mentionné dans la figure (2.7), la phase de détection se décompose en trois étapes: évaluation des résidus, choix du seuil et prise de décision. Ces étapes sont décrites ci-dessous :

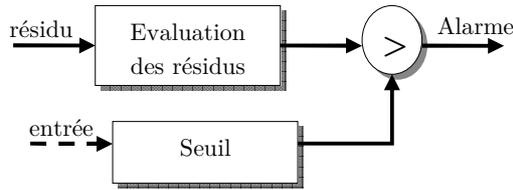


Figure 2.7. Phase de détection.

- *Evaluation des résidus*: en se basant sur le type du modèle à diagnostiquer, deux stratégies se posent. Pour les systèmes déterministes, des stratégies d'évaluation de la norme de résidus sont préférées, et dans le cas des systèmes stochastiques, des méthodes statistiques sont adoptées. Divers fonctions d'évaluations sont proposées dans la littérature. Une étude détaillée est proposée par [Ding, 2008].

Dans des conditions déterministes, la norme  $L_2$  est la plus utilisée pour l'évaluation des résidus. Elle est définie par :

$$\|r\|_2 = \sqrt{\int_0^{\infty} r^T r dt}$$

L'implémentation de la norme  $L_2$  n'est pas faisable, puisque la valeur de  $\|r\|_2$  n'est pas connu à  $t = \infty$ . Par conséquent, elle est calculée sur un intervalle de temps par :

$$\|r\|_{2,[t-\tau,t]} = \sqrt{\int_{t-\tau}^t r^T r dt}$$

L'utilisation de  $\|r\|_{2,[t-\tau,t]}$  au lieu de la norme  $\|r\|_2$  réduit l'optimalité de la méthode de détection de défauts [Ye et al., 2002].

La RMS (la Moyenne Quadratique) est utilisée aussi pour l'évaluation des résidus

$$\|r\|_{RMS} = \sqrt{\frac{1}{\tau} \int_0^\tau r^T r dt}$$

Une assez large classe de fonction d'évaluation utilisée dans la littérature peut être représentée sous la forme suivante:

$$\mathcal{L}_v r = \left( \int_0^t v(t-\tau) |r|^p dt \right)^{\frac{1}{p}}$$

Où le support de la fonction  $v$  permet de décrire une fenêtre temporelle utilisée pour l'évaluation des résidus, éventuellement avec une certaine pondération. Une fonction d'évaluation avec  $p = 2$  est utilisée, par exemple dans [Zhang et al., 2002].

D'autres fonctions d'évaluation des résidus plus couramment utilisées comprennent: la valeur absolue, la valeur moyenne, etc.

Dans le cadre stochastique, sans être exhaustif, nous pouvons citer le test robuste du maximum de vraisemblance (generalized Likelihood Ratio-GLR) proposé par Willsky [Willsky, 1976]. Plus de détails sur les fonctions d'évaluations stochastiques peuvent être trouvés dans [Ding, 2008 ; Blanke et al., 2006 ; Basseville, 2003].

- *Choix du seuil* : dans le cadre idéal (absence d'incertitude du modèle, pas de bruits), les résidus sont supposés avoir une valeur théorique nulle et non nulle dans le cas contraire. La principale difficulté dans la détection est le choix du seuil des résidus vérifiant les propriétés suivantes :

- $r(t) \geq \varepsilon$ , présence d'un défaut
- $r(t) < \varepsilon$ , absence d'un défaut

Un seuil trop grand risque d'engendrer une non-détection: la situation est supposée normale alors qu'elle ne l'est pas. Au contraire, un seuil trop faible entrainera des fausses alarmes: la situation est supposée anormale alors qu'elle ne l'est pas. Ainsi, se pose le problème de déterminer à partir de quel écart peut-on considérer qu'un défaut est potentiellement détecté. La problématique est donc de trouver un seuil optimal qui constituera le compromis idéal entre un taux de non-détection et un taux de fausse alarme minimum. Il existe deux types de seuil: seuil fixe et seuil adaptatif (figure (2.8))

- Seuil fixe: concernant le seuillage, les premiers travaux ont porté sur l'élaboration des seuils fixes, indépendants du temps et des entrées du système. La méthode consiste à détecter l'instant pour lequel les résidus sortent d'une zone délimitée par des seuils autour d'une valeur objectif. Par exemple, [Walker et Gai, 1979] l'ont défini à l'aide de la théorie de Markov. Une incertitude ou une perturbation extérieure peut rendre inadaptée la détection par rapport à un seuil de valeur constante prédéfini.
- Seuil adaptatif: cette méthode permet d'être robuste contre les incertitudes du modèle et le bruit. L'idée consiste à définir des bornes inférieures et supérieures pour le résidu

en fonction des bornes des incertitudes des paramètres et des bruits de mesures. De nombreuses contributions existent dans ce domaine [Ding et Frank, 1991 ; Sauter et al., 1996].

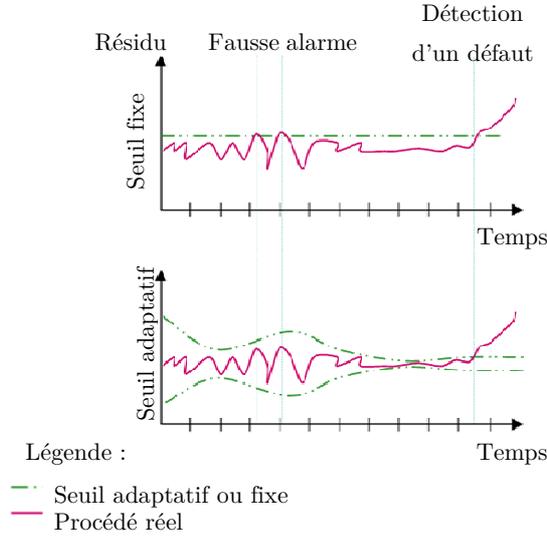


Figure 2.8. Différents types de seuil.

- *La prise de décision*: la troisième étape de l'évaluation d'un résidu est la prise de décision logique. La méthode la plus simple est de faire une comparaison entre le seuil et le résidu, si les résidus dépassent la valeur fixée du seuil, alors une alarme se déclenche indiquant la présence d'un défaut. D'autres méthodes, se basant sur la logique floue ou les réseaux de neurones, servent aussi pour l'évaluation des résidus, pour plus de détails voir [Frank, 1996 (a)].

### c. La localisation

L'objectif de la localisation est de remonter à l'origine du défaut. Elle est nommée encore isolation de défaut. Pour détecter un défaut, un seul résidu est à la limite nécessaire, cependant plusieurs résidus (ou vecteur de résidus) sont souvent requis pour l'isolation des défauts. Pour pouvoir localiser efficacement un ou plusieurs défauts, le vecteur résidu doit avoir un certain nombre de propriétés permettant de caractériser de manière unique chaque défaut.

Soit

$$r_i = g(f_1, f_2, \dots, f_p) \quad i = 1 \dots n$$

Où  $f_i$  est un défaut. Deux types de résidus peuvent se présenter [Patton, 1994]: les résidus directionnels et résidus structurés.

- *les résidus directionnels* : l'idée de cette approche est de générer un résidu sous forme d'un vecteur. En réponse à un défaut donné, le vecteur résidu est orienté suivant une direction privilégiée de l'espace de résidus. Le vecteur de résidu  $r(t)$ , en réaction à un défaut  $f_i(t)$ , s'exploite sous la forme :

$$r(t/f_i) = \gamma_i(t)\rho_i \quad i \in \{1, 2, \dots, n\}$$

Où  $\rho_i$  est un vecteur nommé signature directionnelle du défaut  $i$  dans l'espace des résidus et  $\gamma_i(t)$  est une fonction scalaire qui dépend de l'amplitude de la dynamique du défaut [Gertler, 1998]. Ainsi, la localisation du défaut se réduit à la détermination de la signature la plus proche de la signature réelle obtenue par le calcul des vecteurs de résidus. La figure (2.9) illustre l'exemple d'un problème de localisation de défauts en utilisant des résidus directionnels. Les vecteurs en trait plein représentent les signatures réelles du résidu à des instants différents. Les trois vecteurs en trait pointillé représentent les signatures directionnelles théoriques.

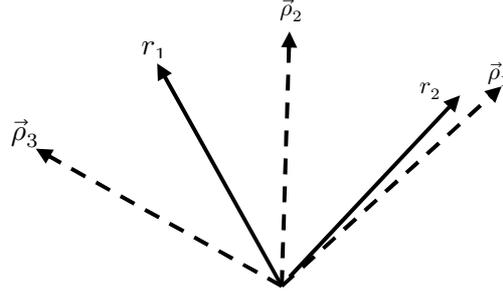


Figure 2.9. Localisation des défauts en utilisant des résidus directionnels

La signature réelle de  $r_2$  est très proche de la signature théorique du défaut 1. Cependant,  $r_2$  ne coïncide pas exactement avec  $\rho_1$  à cause des perturbations. Implicitement, l'influence des perturbations est considérée petite devant celle des défauts. Par contre, le résidu  $r_1$  est plus difficile à évaluer en étant aussi proche de  $\rho_2$  et  $\rho_3$ .

- *Les résidus structurés*: cette approche est amplement utilisée à cause de sa simplicité et son principe de généralité (elle s'applique aux systèmes linéaires et non-linéaires). Les résidus sont conçus de manière à être chacun affectés par un sous ensemble de fautes et robustes par rapport aux fautes restantes. Ainsi, un seul sous-ensemble de résidus réagit, lorsqu'un défaut  $f_i$  apparaît. Par la suite, une table de signature ou table d'incidence regroupe les informations de sensibilité et de robustesse pour les résidus. Elle est définie de la façon suivante:

$$T: R \times F \mapsto [0; 1]$$

$$(r, f) \rightarrow T(r, f)$$

$$\text{Avec } T(r, f) = \begin{cases} 1 & \text{si et seulement si } r \text{ est sensible au défaut } f \\ 0 & \text{si et seulement si } r \text{ est robuste au défaut } f \end{cases}$$

Les dimensions de la table de signature sont déterminées à partir du nombre de capteurs ou actionneurs et du nombre de résidus engendré par la méthode choisie. C'est une table binaire, dont la colonne  $j$  correspond au défaut  $f_j$  et la ligne  $i$  correspond au résidu  $r_i$ .

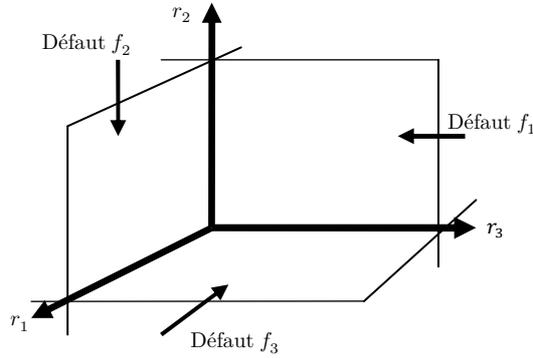


Figure 2.10. Résidus structurés

En employant les termes utilisés par [Gertler, 1992], on qualifie la table de signature de:

- non localisante: si aux moins deux signatures de pannes sont identiques (a).
- faiblement localisante: si toutes les signatures de pannes sont différentes, mais si on modifie une signature en changeant un «1» par «0», on trouve autre signature déjà existante (b).
- fortement localisante: si toutes les signatures de défauts sont différentes et ne peuvent être déduite d'une autre en modifiant un «1» par «0» (c).

$$T_1 = \begin{matrix} & f_1 & f_2 & f_3 & f_4 & f_5 \\ \begin{matrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix} \end{matrix} \quad
 T_2 = \begin{matrix} & f_1 & f_2 & f_3 & f_4 & f_5 \\ \begin{matrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix} \end{matrix} \quad
 T_3 = \begin{matrix} & f_1 & f_2 & f_3 & f_4 & f_5 \\ \begin{matrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{pmatrix} \end{matrix}$$

(a)

(b)

(c)

Figure 2.11. Différents types de tables de signature.

De nombreux schémas sont proposés dans la littérature pour concevoir des résidus structurés, parmi elles, on peut citer:

- Schémas d'observateur dédié (DOS- Dedicated Observer Scheme),
- schémas d'observateur généralisé (GOS- Generalized Observer Scheme).

Dans un schéma d'observateur dédié (DOS), le  $i^{\text{ème}}$  observateur est piloté par la  $i^{\text{ème}}$  entrée et toutes les sorties; les  $m - 1$  autres entrées restantes sont considérées comme inconnues et la sortie de cet  $i^{\text{ème}}$  observateur est insensible aux défauts des entrées non utilisées.

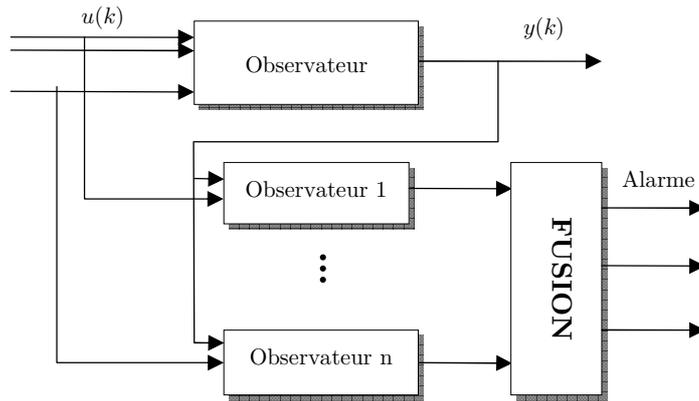


Figure 2.12. Observateur DOS pour la détection des défauts actionneurs.

La figure (2.13) est un schéma d'observateur généralisé (GOS). Elle est relative à une structure où le  $i^{\text{ème}}$  observateur est piloté par toutes les entrées sauf la  $i^{\text{ème}}$  et toutes les sorties. La sortie de cet observateur est donc sensible aux défauts de toutes les entrées sauf ceux de l'entrée  $i$ . Le principe est le même pour la reconstruction des défauts capteurs.

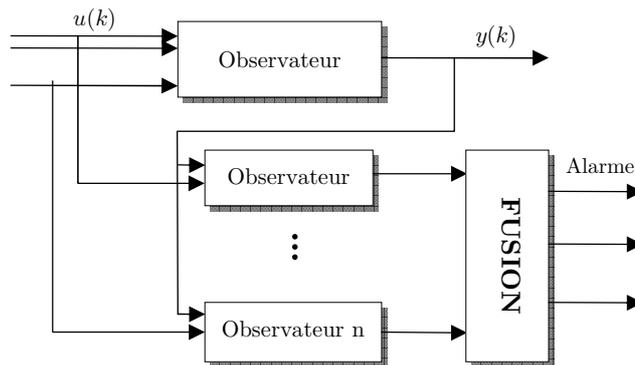


Figure 2.13. Observateur GOS pour la détection des défauts capteurs.

#### d. L'identification

Bien qu'elle soit très importante dans un schéma de diagnostic, l'identification ou l'estimation de défaut a reçu moins d'intérêt dans les recherches scientifiques et spécialement pour les systèmes non-linéaires. L'objectif de cette procédure consiste à déterminer (identifier) la caractéristique précise de la défaillance. L'identification ou l'estimation du défaut est une tâche plus délicate et qui nécessite d'utiliser un modèle de comportement du système de niveau élevé de connaissance sur les défaillances (le graphe du défaut, sa structure, sa dynamique). Obtenir une estimation du défaut permet bien entendu de donner une image beaucoup plus précise de l'état du système et de mettre en œuvre des procédures de commande à savoir des procédures tolérantes aux défaillances. Cette partie sera expliquée plus en détails dans le chapitre 4.

#### 2.2.4. Caractéristiques souhaitables d'un système de diagnostic

A partir de la classification des méthodes de diagnostic du paragraphe (2.2.2) de ce chapitre, nous avons mis en avant qu'il est très difficile de comparer ces méthodes. Afin de définir une comparaison, il est utile d'identifier un ensemble de caractéristiques souhaitables qu'un système de diagnostic devrait avoir. Ainsi les différentes approches peuvent être évaluées à partir de cet ensemble commun de conditions.

- *La rapidité de détection*: c'est la capacité d'un système de diagnostic à détecter et à isoler rapidement une défaillance. La réponse rapide d'un système de diagnostic est de détecter les défauts difficiles à être détecter (graduels, intermittents). Détecter des défauts avant leurs complètes apparitions est très bénéfique pour la sûreté de fonctionnement. Cependant, le système de diagnostic sera sensible aux fausses alarmes.
- *L'isolabilité*: c'est la capacité de localisation d'origine du défaut. Le degré d'isolabilité est lié à la structure du système de diagnostic et aussi à la méthode de conception des résidus. Les incertitudes de modélisation et les perturbations limitent les caractéristiques d'isolabilité d'un système de diagnostic. C'est pour cette raison qu'il y a un compromis entre l'isolabilité et le rejet des incertitudes.
- *L'identifiabilité d'un défaut*: c'est la détermination du type et de la nature du défaut. Obtenir une bonne identification du défaut est difficile à réaliser à cause de la présence des perturbations et des incertitudes de modélisation.
- *La robustesse*: la robustesse peut être définie comme étant le rapport entre une sensibilité maximale vis-à-vis du défaut recherché et une sensibilité minimale vis-à-vis des autres perturbations (bruits de mesures, variations paramétriques, incertitudes de modélisation...)

Une amélioration de la robustesse peut être réalisée par le choix approprié :

- de la méthode de génération des résidus qui dépend des défauts à détecter et des perturbations,
  - des fonctions de prise de décisions en fonction du type de résidus générés (structurés ou directionnels).
- *L'adaptabilité*: les conditions opérationnelles d'un système changent en raison des modifications des entrées extérieures ou des modifications des conditions environnementales (changement de la qualité des matières premières). Par conséquent, un système de diagnostic s'adapte à ses changements dans le but de maintenir les performances de diagnostic.
  - *L'identification des défauts multiples*: la capacité d'identifier des défauts et de classer correctement les défauts multiples qui peuvent coexister dans un système. C'est un problème du à l'interaction de la plupart des défauts. Une autre raison est le nombre de combinaisons différentes de défauts multiples à prendre en compte, ce qui engendre une combinatoire prohibitive pour de grands nombres de processus.

Donc nous pouvons récapituler que ses performances consistent à minimiser (retards à la détection, taux de fausse alarme et de mauvaise détection, et le temps de calcul pour une utilisation en temps réel) et à maximiser (la sensibilité à des défauts graduels et l'insensibilité aux bruits et aux incertitudes sur les paramètres du modèle).

Il est clair que ces critères sont contradictoires, ainsi une optimisation simultanée de ces critères semble difficile. Les méthodes de détection doivent être choisies selon un compromis précisé dans un cahier de charge, en avantageant certains critères et en pénalisant d'autres.

Jusqu'à ce point, nous avons donné une description des méthodes de diagnostic. Aussi, nous nous sommes longuement arrêtés sur les méthodes de diagnostic à base de modèle et les différentes étapes suivies pour le diagnostic: la génération de résidu, la détection, la localisation et l'identification des défauts. À la fin de cette première partie, une présentation des performances des méthodes de diagnostic qui vise principalement à minimiser la sensibilité d'un système de diagnostic aux perturbations et aux incertitudes et à maximiser sa sensibilité aux défauts, a été donnée.

Nous allons maintenant étudier plus en détails l'observabilité des systèmes non linéaires.

### **2.3. Observabilité et observateurs grand gain**

Bien qu'elles soient indispensables pour la description mathématique de la dynamique du système, il arrive souvent que certaines variables d'état ne soient pas accessibles à la mesure.

Pour des raisons techniques et économiques, ce problème est particulièrement prépondérant dans le domaine d'épuration des eaux usées où l'instrumentation des mesures est limitée par le manque de capteurs fiables et peu coûteux.

L'accès en ligne de l'état du système constitue un élément essentiel pour la conduite automatisée, la commande ou encore le diagnostic des pannes. L'état du système est composé par un ensemble de variables décrivant l'évolution du processus. Ces variables décrivent généralement un comportement physique et peuvent inclure certains paramètres intervenant dans le modèle. Malheureusement, dans la plupart des procédés physiques seulement quelques mesures en ligne donnant une information sur l'évolution du système sont à la disposition de l'opérateur. Un des problèmes auxquels sont confrontés les industriels est le manque d'instrumentation en ligne. En particulier, les mesures courantes dans les stations d'épurations biologiques sont réalisées le plus souvent hors ligne et nécessitent des procédures d'analyse engendrant des retards importants.

Une alternative pour pallier à ces limitations consiste à développer des capteurs logiciels, ou encore appelés observateurs d'état ou estimateurs en termes de l'automatique. Ces outils permettent de reconstruire une estimation des variables non mesurées ou des paramètres mal connus, à partir des mesures.

### 2.3.1. Généralité sur les observateurs

La figure (2.14) fournit le schéma de principe d'un observateur.

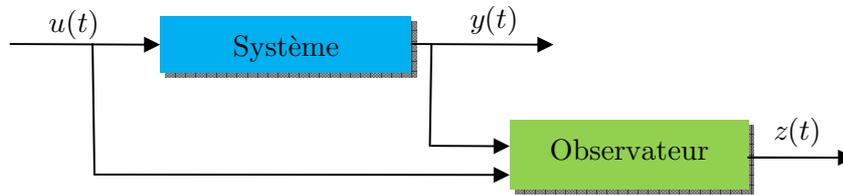


Figure 2.14. Schéma de principe d'un observateur

#### Définition 2.1.

Un observateur est un système dynamique auxiliaire dont les entrées sont constituées des vecteurs d'entrées  $u(t)$  et de sortie  $y(t)$  du système. Il a pour rôle de donner une estimation  $\hat{x}(t)$  du vecteur d'état  $x(t)$  du système considéré. Cet état estimé sera utilisé à des fins de commande du système ou de détection et localisation des défauts. Un tel système peut être représenté d'une façon générale comme suit :

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

$$\hat{x}(t) = g(\hat{x}(t), u(t), y(t), t)$$

Tel que

$$(i) \quad \hat{x}(0) = x(0) \Rightarrow \hat{x}(t) = x(t), \forall t \geq 0;$$

$$(ii) \quad \|\hat{x}(t) - x(t)\| \rightarrow 0 \text{ lorsque } t \rightarrow \infty;$$

Si (ii) est vérifié pour tout  $x(0)$  et  $\hat{x}(0)$ , l'observateur est global.

Si (ii) est vérifié avec une convergence exponentielle, l'observateur est exponentiel.

Si (ii) est vérifié avec une vitesse de convergence à régler, l'observateur est réglable.

La conception et l'étude d'un observateur peuvent être divisées en trois étapes :

- *L'observabilité* : c'est une étape primordiale pour la mise au point d'un observateur. Elle permet de savoir si la reconstitution de l'état à partir des mesures est réalisable.
- *La convergence* : c'est l'étude de la convergence de l'état estimé vers l'état réel, et à quelle vitesse (convergence exponentielle, convergence asymptotique,...) ?
- *La stabilité* : la stabilité d'un algorithme de contrôle qui s'appuie sur l'estimation fournie par l'observateur.

### 2.3.2. Système en considération

On considère un système composé de deux équations différentielles de la forme :

$$\Sigma \begin{cases} \dot{x}(t) = f(x(t), u(t), t) \\ y(t) = h(x(t), u(t), t) \\ x(0) = x_0 \end{cases} \quad (2.19)$$

Où

- $x(t) \in X$  est l'état du système.  $X$  est une variété de dimension  $n$ ,
- $u(t) \in U_{adm} \subset \mathbb{R}^{n_u}$  est un ensemble des valeurs de contrôles dites admissibles,
- $y(t)$  est l'ensemble des fonctions mesurables (ou sorties) à valeurs dans  $\mathbb{R}^{n_y}$ ,
- $f$  est une famille de champs de vecteurs paramétrée par  $u$ , et
- $h$  est une application de  $X \times \mathcal{U}_{adm}$  à valeurs dans  $\mathbb{R}^{n_y}$ .

**Les entrées** sont des fonctions définies sur un intervalle semi-ouvert de la forme  $[0, T[$  (avec la possibilité que  $T = +\infty$ ). Les fonctions  $u(\cdot)$  sont des fonctions mesurables et bornées sur tout sous-intervalle de  $[0, T[$ . L'ensemble de fonction correspondante est  $L^\infty(\mathcal{U}_{adm})$ .

**Les sorties** aussi sont des fonctions définies sur un intervalle semi-ouvert de la forme  $[0, T(u)[$ . Les fonctions de sorties sont des fonctions mesurables et bornées sur tout sous-intervalle de  $[0, T(u)[$ . L'ensemble de fonction correspondante est  $L^\infty(\mathbb{R}^{n_y})$ .

L'ensemble des systèmes ayant la forme  $(\Sigma)$  est noté  $\mathcal{S} = \{\Sigma = (f, h)\}$ . La propriété de généricité ou non-généricité des systèmes observables est considérée par rapport à l'ensemble  $\mathcal{S}$ .

### 2.3.3. Observabilité des systèmes non-linéaires

Avant de construire un observateur, on doit se demander dans quelles conditions ceci est possible ?

Considérons un système linéaire défini par les équations suivantes :

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases} \quad (2.20)$$

Avec  $x \in X$ ,  $u \in \mathbb{R}^{n_u}$  et  $y \in \mathbb{R}^{n_y}$ ,  $x$  représente toujours l'état du système,  $u$  l'entrée et  $y$  la sortie. Sous des conditions d'observabilité bien connues :

$$\text{rang} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n, \text{ le système (2.20) est observable.}$$

Dans le cas linéaire, cette condition est indépendante de l'entrée (la condition du rang dépend que de  $A$  et  $C$ ), et elle est également suffisante pour garantir l'existence d'un observateur à vitesse de convergence exponentielle et arbitrairement rapide [Luenberger, 1971; Aeyels, 1981].

Pour les systèmes non-linéaires, malgré l'abondance des résultats disponibles, il n'existe pas une solution systématique. Une des raisons en est que, contrairement au cas linéaire, l'observabilité d'un système non-linéaire dépend de l'entrée appliquée.

On peut voir facilement à travers l'exemple suivant : Soit le système non-linéaire suivant:

$$\begin{cases} \dot{x}_1 = x_2 u \\ \dot{x}_2 = 0 \\ y = x_1 \end{cases} \quad (2.21)$$

La variable d'état  $x_2$  peut être exprimée en fonction de la sortie et de l'entrée,  $x_2 = \frac{\dot{y}}{u}$ . Il est alors évident que  $x_2$  peut être connu à partir des grandeurs disponibles (sortie et entrée) si  $u \neq 0$ , l'observabilité du système (2.21) dépend donc de l'entrée.

Il existe dans la littérature plusieurs définitions de la notion d'observabilité introduites surtout par Sontag [Sontag, 1979], par Herman et Krener [Herman et Krener, 1977], et Gauthier et Kupka dans [Gauthier et Kupka, 2001]. Nous donnons au début la définition classique de l'observabilité. Nous nous focalisons par la suite sur la notion d'observabilité introduite dans les travaux [Gauthier et al., 1991], [Gauthier et Kupka, 1996], [Jouan et Gauthier, 1996].

Considérons le système non-linéaire (2.19), la définition classique d'observabilité est liée à l'indistinguabilité. Notons par  $x_u(t, x_0, t_0)$ , la trajectoire à l'instant  $t$ , solution du système (2.19) issue de la condition initiale  $x_0 = x(t_0)$  ( $t_0 \geq 0$ ) et des commandes appliquées  $u(t)$ , pour tout  $t \geq t_0$ .

### Définition 2.2

Soit le système non-linéaire (2.19). Soient deux états initiaux  $x_0^1$  et  $x_0^2$  distincts. Ils sont dits indiscernables si pour toute fonction d'entrée  $u(t)$  et pour tout  $t \geq t_0 \geq 0$ , les sorties  $h(x_u(t, x_0^1, t_0))$  et  $h(x_u(t, x_0^2, t_0))$  qui en résultent sont égales.

C'est-à-dire qu'à partir de deux conditions initiales différentes, la sortie du système est la même dans les deux cas pour une même commande appliquée. Ainsi,

### Définition 2.3

Le système (2.19) est dit observable s'il ne possède pas de couple d'états initiaux distincts indiscernables.

Gauthier et Kupka ont introduit de nouvelles notions d'observabilité. Pour une entrée non nulle, connue, il y a assez d'information dans la sortie pour reconstruire l'état initial. Une entrée  $u$  étant fixée,  $u \neq 0$ , l'application :

[État initial  $\rightarrow$  Sortie en tant que fonction] est injective.

### Définition 2.4

Pour le système  $\Sigma$ , on définit l'application entrées / sorties par

$$\begin{aligned} PX_{\Sigma, u}: X &\rightarrow L^\infty(\mathbb{R}^{n_y}) \\ x_0 &\mapsto y(\cdot) \end{aligned}$$

Le système  $\Sigma$  est uniformément observable (ou juste observable), si pour tout entrée  $u(\cdot) \in \mathcal{C}$ ,  $PX_{\Sigma, u}$  est injective.

Malheureusement cette notion d'observabilité devient instable suite à une faible variation du système. Exemple  $(x \mapsto x^3)$  est injective, mais pour tout  $\epsilon > 0$ ,  $(x \mapsto x^3 - \epsilon x)$  n'est plus injective. Pour cette raison, Gauthier et Kupka ont ajouté la condition d'immersion (injectivité infinitésimale) en plus de l'injectivité.

### Définition 2.5

- Considérons la variation d'ordre 1 du système  $(\Sigma)$  :

$$TX_{\Sigma} \begin{cases} \dot{x}(t) = f(x, u) \\ \dot{\xi}(t) = D_x f(x, u)\xi \\ \hat{y} = d_x h(x, u)\xi \end{cases} \quad (2.22)$$

Où

- (i)  $(x, \xi) \in TX$  est l'état de  $(TX_{\Sigma})$ ,
  - (ii)  $d_x h$  est le différentiel de  $h$  par rapport à  $x$ , et
  - (iii)  $D_x f$  est l'application tangente de  $f$  (représenté par les matrices jacobéenne de  $h$  et  $f$  par rapport à  $x$ )
- L'application entrées / sorties de  $TX_{\Sigma}$  est notée  $PTX_{\Sigma, u}$ . En effet, cette application est le différentiel de  $PX_{\Sigma, u}$  par rapport à  $x_0$ .

### Définition 2.6

Le système est dit uniformément infinitésimalement observable, si pour chaque  $u(\cdot) \in C$ , et pour chaque  $x_0 \in X$  ( $x_0$  paramétré), toutes les applications tangentes  $PTX_{\Sigma, u}|_{x_0}$  sont injectives.

Puisque l'application entrées/sorties considérée dans cette définition est linéaire, alors la propriété d'injectivité est devenue stable (topologiquement).

Finalement, une troisième définition de l'observabilité a été proposée en utilisant la notion de jets d'ordre  $k$ . Les jets d'ordre  $k$  pour une fonction lisse  $u$  en un point  $t = 0$  sont définis par :

$$j^k u = (u(0), \dot{u}(0), \dots, u^{k-1}(0))$$

### Définition 2.7

- $\Sigma$  est dit différentiellement observable d'ordre  $k$  si l'application  $\Phi_k^{\Sigma}$  suivante est injective.

$$\begin{aligned} \Phi_k^{\Sigma}: X &\rightarrow \mathbb{R}^{kn_y} \\ x_0 &\mapsto j^k y \end{aligned}$$

- $\Sigma$  est dit fortement différentiellement observable d'ordre  $k$  si l'application  $\Phi_{k, j^k u}^{\Sigma}$  suivante est une immersion injective.

$$\begin{aligned} \Phi_{k, j^k u}^{\Sigma}: X &\rightarrow \mathbb{R}^{kn_y} \\ x_0 &\mapsto j^k y \end{aligned}$$

A partir de la définition (2.7), l'observabilité différentielle forte implique l'observabilité différentielle. Aussi, un autre résultat a été établi : l'observabilité différentielle (forte) dans la classe des fonctions  $\mathcal{C}^\infty$  implique l'observabilité dans la classe des fonctions  $\mathcal{C}^\infty$ .

***Théorème 2.1 [Gauthier et Kupka, 2001]***

*Pour un système analytique  $(\Sigma)$  les propriétés suivantes sont équivalentes :*

1.  $\Sigma$  est observable dans la classe des fonctions analytiques
2.  $\Sigma$  est observable dans la classe des fonctions  $L^\infty$

Cela signifie que pour un système analytique, l'observabilité différentielle forte et l'observabilité différentielle implique l'observabilité dans la classe des fonctions  $L^\infty$ .

Une fois que l'étude de l'observabilité aboutit à un système observable, une transformation de coordonnées sera nécessaire pour mettre le système sous la forme normale d'observabilité. Un rappel de ces différentes formes normales fera l'objet de la section suivante.

**2.3.4. Formes normales d'observabilité**

L'idée principale de cette théorie consiste en l'existence de deux situations distinctes, qui dépendent du nombre d'entrées par rapport à celui des sorties. Soit l'observabilité est une propriété générique, soit elle ne l'est pas. Une explication de ces deux situations particulières est donnée dans les deux sous-sections ci-dessous.

**a. Cas où  $n_y > n_u$**

Le premier cas se produit lorsque le nombre de sorties est supérieur au nombre d'entrées, soit  $n_y > n_u$ . Cette situation se définit par deux théorèmes ; le premier énonce la propriété de généricité. Le deuxième introduit la forme canonique d'observabilité correspondante.

***Théorème 2.2 [Gauthier et Kupka, 2001]***

- *L'ensemble des systèmes fortement différentiellement observable d'ordre  $2n + 1$  est résiduel dans  $\mathcal{S}$ .*
- *L'ensemble des systèmes analytiques fortement différentiellement observable d'ordre  $2n + 1$  qui est en plus observable dans la classe des fonctions  $L^\infty$  est dense dans  $\mathcal{S}$ .*

***Théorème 2.3 [Busvelle et Gauthier, 2003]***

*La propriété suivante est générique dans  $\mathcal{S}$  : Soit  $k = 2n + 1$ . Pour tout compact  $\Gamma \subset X$  et  $u, \dot{u}, \dots, u^k$  bornée, alors les applications*

$$\Phi_{k,j^k u}^\Sigma: X \rightarrow \mathbb{R}^{kn_y}$$

$$x(t) \mapsto (y(t), \dot{y}(t), \dots, y^{k-1}(t))$$

*sont des immersions injectives qui transforment les trajectoires du système  $\Sigma$  (restreint à  $\Gamma$ ) aux trajectoires du système suivant:*

$$\begin{cases} y = z_1 \\ \dot{z}_1 = z_2 \\ \vdots \\ \dot{z}_{k-1} = z_k \\ \dot{z}_k = \phi_k(z_1, z_2, \dots, z_k, u, \dot{u}, \dots, u^{(k)}) \end{cases} \quad (2.23)$$

Puisque toutes les applications  $\Phi_{k,j}^{\Sigma}$  sont des immersions injectives restreint à  $\Gamma$  (théorème 2.3), alors le système de la forme (2.23) est observable, fortement observable, différentiellement observable. C'est-à-dire que l'ensemble du système qui peut être mis sous la forme (2.23) est contenu dans l'ensemble des fonctions observables.

### b. Cas où $n_y \leq n_u$

Dans la section précédente, nous avons représenté l'observabilité comme étant une propriété générique dans le cas où le nombre de sorties est supérieur à celui des entrées ( $n_y > n_u$ ). Cette propriété n'est plus valable lorsque ( $n_y \leq n_u$ ). Nous limitons l'exposition au cas des systèmes analytiques mono-sortie, à savoir  $n_y = 1$ . L'étude de l'observabilité est faite dans ce cas, à l'aide d'un outil appelé drapeau canonique pour une distribution.

Comme nous le verrons dans les théorèmes suivants, l'observabilité dépend de l'existence ou l'inexistence du drapeau canonique pour une distribution.

#### Définition 2.8

- Considérons un système  $\Sigma = (f, h) \in \mathcal{S}$ . Le drapeau canonique de distributions  $D(u)$  est défini par :

$$\begin{cases} D(u) = \{D^0(u) \supset D^1(u) \supset \dots \supset D^{n-1}(u)\} \\ D^0(u) = \text{Ker}(d_x h) \\ D^{k+1}(u) = D^k(u) \cap \text{Ker}(d_x L_f^{k+1} h) \end{cases} \quad (2.24)$$

Où  $L_f h$  est la dérivée de Lie de  $h$  dans la direction de  $f$ .  $\text{Ker}$  désigne le noyau d'une application. Le contrôle  $u(t)$  est considéré fixe.

- Si les distributions  $D^i(u)$  sont de rang constant  $n - i - 1$ , et sont indépendantes de  $u(\cdot)$ , alors  $D(u)$  est dit drapeau canonique uniforme.

La propriété

$$(\Sigma \text{ admet un drapeau canonique uniforme})$$

est non-générique. La non-généricité de l'observabilité est donnée par les deux théorèmes suivants.

#### Théorème 2.4 [Busse et How, 2002]

Le système  $(\Sigma)$  admet un drapeau canonique uniforme si et seulement si pour tout  $x^0 \in X$ , il existe un voisinage de  $x^0$  de coordonnée  $(V_{x^0}, x)$  dans lequel la restriction de  $(\Sigma)$  à  $V_{x^0}$  s'écrit:

$$\begin{cases} y = h(x_1, u) \\ \dot{x}_1 = f_1(x_1, x_2) \\ \dot{x}_2 = f_2(x_1, x_2, x_3) \\ \vdots \\ \dot{x}_{n-1} = f_{n-1}(x_1, x_2, \dots, x_n, u) \\ \dot{x}_n = f_n(x_1, x_2, \dots, x_n, u) \end{cases} \quad (2.25)$$

Où  $\frac{\partial h}{\partial x_1}$  et  $\frac{\partial f_i}{\partial x_{i+1}}$ ,  $i = 1, \dots, n-1$ , ne s'annulent jamais sur  $V_{x^0} \times \mathcal{U}_{adm}$ .

Ainsi un système de la forme (2.25) est infinitésimalement observable, observable et différentiellement observable d'ordre  $n$ . Par conséquent, si un système admet un drapeau canonique uniforme sur  $V_{x^0} \times \mathcal{U}_{adm}$ , il a toutes les propriétés d'observabilités. A ce stade, les relations obtenues peuvent être résumées par le schéma d'équivalence suivant.

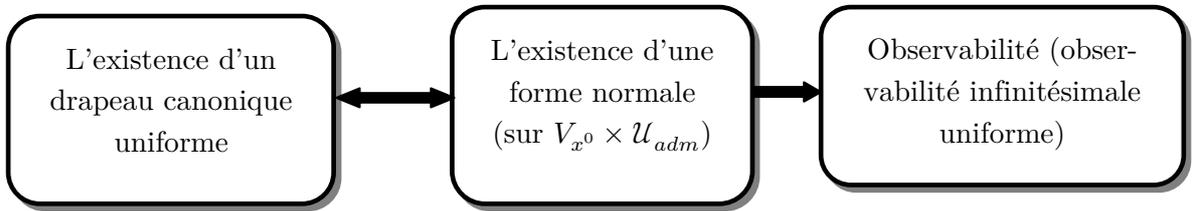


Figure 2.15. Le diagramme d'équivalence de l'observabilité

Le théorème suivant permet de compléter le diagramme de l'observabilité ci-dessus et préciser la relation d'équivalence entre l'observabilité infinitésimale et l'existence de la forme normale (2.25).

***Théorème 2.5 [Busse et How, 2002]***

*Si  $\Sigma$  est uniformément infinitésimalement observable, alors, sur le complémentaire d'un ensemble sous-analytique de  $X$ , de codimension 1,  $(\Sigma)$  admet un drapeau canonique uniforme.*

Pour les systèmes affines en contrôle de la forme :

$$\begin{cases} \dot{x} = f(x) + \sum_{i=1}^p g_i(x)u_i \\ y = h(x) \end{cases} \quad (2.26)$$

Nous définissons ensuite la fonction

$$\begin{aligned} \Phi: X &\rightarrow \mathbb{R}^n \\ x &\mapsto (h(x), L_f h(x), \dots, L_f^{n-1} h(x)) \end{aligned}$$

***Lemme 2.1***

$\Sigma$  est observable  $\Rightarrow \Phi$  est de rang maximum dans un ouvert dense  $V$  de  $X$ .

Soit un sous-ensemble  $W \subset X$  tel que la restriction de  $\Phi$  à  $W$  est un difféomorphisme. Ainsi les théorèmes ci-dessus deviennent :

**Théorème 2.6 [Busse et How, 2002]**

Supposons que  $\Sigma$  est observable. Alors la restriction de  $\Phi|_W$  transforme  $\Sigma$  en un système de la forme:

$$\begin{aligned}
y &= x_1 \\
\dot{x}_1 &= x_2 + \sum_{i=1}^p g_{1,i}(x_1)u_i \\
\dot{x}_2 &= x_3 + \sum_{i=1}^p g_{2,i}(x_1, x_2)u_i \\
&\vdots \\
\dot{x}_2 &= x_n + \sum_{i=1}^p g_{n-1,i}(x_1, x_2, \dots, x_n)u_i \\
\dot{x}_n &= \psi(x) + \sum_{i=1}^p g_{n-1,i}(x_1, x_2, \dots, x_{n-1}, x_n)u_i
\end{aligned} \tag{2.27}$$

Inversement si un système est sous la forme (2.19), sur un ouvert  $\Omega \subset \mathbb{R}^n$  alors il est observable. Finalement, nous soulignons les trois formes canoniques (2.23), (2.25) et (2.27) puisqu'elles traduisent l'observabilité.

**2.3.5. Formes normales Multi-entrées Mono-sortie (MISO)**

La forme normale introduite dans cette section est une forme Multi-entrées Mono-sortie. Nous utilisons les systèmes MISO pour la simplicité et la clarté de l'exposé. En effet, dans le cas des systèmes Multi-entrées et Multi-Sorties (MIMO), il n'existe pas de forme normale unique, donc la définition de l'observateur doit être modifiée en fonction de chaque cas spécifique. Dans le chapitre suivant, une application d'un observateur grand gain adaptatif à l'estimation d'état d'une station d'épuration biologique à boues activées qui a une forme MIMO sera développée. Le système considéré est composé de deux équations différentielles de la forme :

$$\begin{cases} \dot{x}(t) = A(u)x + b(x, u) \\ y = C(u)x \end{cases} \tag{2.28}$$

Avec

- $x(t) \in \mathcal{X} \subset \mathbb{R}^n$ , où  $\mathcal{X}$  est compact,
- $y(t) \in \mathbb{R}$ ,
- $u(t) \in \mathcal{U}_{adm} \subset \mathbb{R}^{n_u}$ , est borné.

Les matrices  $A(u)$  et  $C(u)$  sont définies par :

$$A(u) = \begin{pmatrix} 0 & a_2(u) & 0 & \dots & 0 \\ & 0 & a_3(u) & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 0 & a_n(u) \\ & & & & 0 \end{pmatrix}$$

$$C(u) = (a_1(u) \quad 0 \quad \dots \quad 0)$$

Avec  $0 < a_m \leq a_i(u) \leq a_M$  pour tout  $u(t)$  dans  $\mathcal{U}_{adm}$  et  $i = 1, \dots, n$ . Nous supposons que  $b(x, u)$  est un champ de vecteur à support compact ayant la structure triangulaire suivante :

$$b(x, u) = \begin{pmatrix} b_1(x_1, u) \\ b_2(x_1, x_2, u) \\ \vdots \\ b_n(x_1, \dots, x_n, u) \end{pmatrix}$$

$b^*(x, u)$  est la matrice jacobéenne de  $b(x, u)$ . Le champ de vecteurs  $b(x, u)$  est supposé Lipchitz.

Dans les sections précédentes, nous avons rappelé les notions d'observabilité des systèmes non-linéaires déterministes élaborées par Gauthier et Kupka [Gauthier et Kupka, 2001]. Une présentation des deux principaux types d'observateurs à grand gain et de leurs théorèmes de convergence respectifs fera l'objet de la section suivante.

### 2.3.6. Observateur à grand-gain

Les premières descriptions d'un observateur à grand-gain peuvent être trouvées dans les références suivantes [Gauthier et al., 1992 ; Deza et al., 1992 (a) ; 1992 (b) ; Tornambé, 1992 ; Esfandiari et Khalil, 1992]. Le premier type d'observateur à grand gain est celui de Luenberger [Gauthier et al., 1992]. Le second est l'observateur de Kalman étendu grand gain. Cet observateur a une structure similaire à celui au filtre de Kalman étendu standard [Gelb, 1974 ; Grewal et al., 2007]. Ce dernier est adapté à la forme canonique d'observabilité en utilisant un scalaire (noté  $\theta$ ).

#### a. Observateur à grand gain de type Luenberger

##### *Définition 2.9*

Supposons que tous les coefficients  $a_i(u)$  de la forme normale (2.28) soient égaux à 1. Les observateurs à grand gain de type de Luenberger s'écrivent sous la forme :

$$\dot{z}(t) = A z + b(z, u) - K_\theta (Cz - y(t)) \quad (2.29)$$

Où  $K_\theta = \Delta K$ ,  $\Delta$  est une matrice diagonale dont les coefficients diagonaux sont  $(\Delta)_{j,j} = \theta^j$  où  $\theta$  est un paramètre supposé grand et  $K$  est une matrice dont les coefficients sont choisis tel que la matrice  $(A - KC)$  soit Hurwitz.

La convergence de l'observateur de type de Luenberger est exprimée par le théorème suivant :

##### *Théorème 2.7*

Pour tout  $a > 0$ , il existe un  $\theta$  assez grand de sorte que  $\forall (x_0, z_0) \in (\chi, \chi)$ , il existe un polynôme  $k$  de degré  $n$ , tel que

$$\|z(t) - x(t)\|^2 \leq k(a)e^{-at} \|z_0 - x_0\|^2$$

Ce théorème exprime que les observateurs à grand-gain de type de Luenberger sont des observateurs exponentiels. Il est à signaler que l'utilisation d'un observateur à grand gain peut conduire à une grande sensibilité au bruit de l'observation, car il amplifie l'erreur d'observation. D'où le compromis classique lors du choix de  $\theta$  entre performance et robustesse. Un grand  $\theta$  assure une convergence rapide mais amplifie le bruit.

### ***b. Observateur de Kalman étendu grand gain***

Le filtre de Kalman est un des outils les plus utilisés pour l'estimation d'état et des paramètres des systèmes linéaires [Kalman, 1960 ; Kalman et Bucy, 1961]. Il estime de façon optimale l'état du système linéaire. C'est un estimateur itératif basé sur une réaction de récurrence, ce qui signifie que seul l'état précédemment estimé et les mesures actuelles sont nécessaires pour le calcul de l'estimation d'état courant.

Le filtre de Kalman étendu est l'extension de ce filtre aux systèmes non-linéaires. La linéarisation du système, initialement non-linéaire, s'effectue autour de la trajectoire estimée par le filtre.

Les équations d'état et d'observation d'un filtre de Kalman étendu s'écrivent alors

#### ***Définition 2.10***

$$\begin{cases} \dot{z}(t) = A(u)z + b(z, u) - S^{-1}C'R^{-1}(Cz - y) \\ \dot{S}(t) = -(A(u) + b^*(z, u))'S - S(A(u) + b^*(z, u)) + C'R^{-1}C - SQS \end{cases} \quad (2.30)$$

Où les matrices  $Q$  et  $R$  sont comme suit :

- Dans le cas stochastique,  $Q$  et  $R$  sont respectivement les matrices de covariance de bruit de l'état et de mesure.
- Dans le cas déterministe,  $Q$  et  $R$  sont des matrices symétriques définies positives qui n'ont pas d'interprétations stochastiques mais qui peuvent être interprétés comme matrices de coûts.

Le filtre de Kalman étendu a été appliqué fréquemment et avec succès. Il a été démontré dans [Dubrovin et al., 1984] que cet observateur est localement exponentiellement convergent i.e. il converge exponentiellement pour  $x_0$  assez proche de  $z_0$  (dans le cadre stochastique, [Hockett, 1990]). En dépit de la popularité du filtre de Kalman étendu, le caractère purement local le rend impropre aux procédés très non-linéaires et ceux qui sont soumis à de grandes perturbations non modélisées et non mesurées. Un autre cas pour lequel le filtre de Kalman étendu ne fonctionne pas correctement est le cas des systèmes non-linéaires dont on ne connaît pas bien l'état initial.

Ainsi, d'un point de vue pratique, le filtre de Kalman étendu présente de bonnes performances. Cependant d'un point de vue théorique, il ne repose sur aucun théorème. D'après l'article [Busvelle et Gauthier, 2002], le filtre de Kalman étendu lorsqu'il est appliqué à un système mis sous une forme canonique d'observabilité alors il acquiert des propriétés de convergence globales.

#### ***Définition 2.11***

$$\begin{cases} \dot{z}(t) = A(u)z + b(z, u) - S^{-1}C'R^{-1}(Cz - y) \\ \dot{S}(t) = -(A(u) + b^*(z, u))'S - S(A(u) + b^*(z, u)) + C'R^{-1}C - SQ_\theta S \end{cases} \quad (2.31)$$

$R$  est une matrice définie positive.  $Q_\theta$  est définie par  $Q_\theta = \theta^2 \Delta^{-1} Q \Delta^{-1}$  où  $\theta > 1$  est un paramètre fixe et

$$\Delta = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{1}{\theta} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & \frac{1}{\theta^{n-1}} \end{pmatrix}$$

Les matrices jacobiennes de  $f$  et  $h$  (calculées par rapport à  $x$ ) sont utilisées dans l'équation de Riccati suivante:

$$\dot{S}(t) = - \left( \frac{\partial f}{\partial x} \Big|_{x=z} \right)' S - S \left( \frac{\partial f}{\partial x} \Big|_{x=z} \right) + \left( \frac{\partial h}{\partial x} \Big|_{x=z} \right)' R^{-1} \left( \frac{\partial h}{\partial x} \Big|_{x=z} \right) - S Q_{\theta} S$$

**Théorème 2.8** [Deza et al., 1992 ; Gauthier et Kupka, 2001]

Pour  $\theta$  assez grand et pour tout  $T > 0$ , l'observateur de Kalman étendu grand gain (2.31), satisfait l'inégalité ci-dessous pour tout  $t > \frac{T}{\theta}$  :

$$\|z(t) - x(t)\|^2 \leq \theta^{n-1} k(T) \left\| z \left( \frac{T}{\theta} \right) - x \left( \frac{T}{\theta} \right) \right\|^2 e^{-(\theta w(t) - \mu(t)) \left( t - \frac{T}{\theta} \right)}$$

Où  $k(T)$ ,  $w(t)$  et  $\mu(t)$  sont des fonctions continues et positives.

### 2.3.7. Observateurs à grand-gain adaptatifs

Au début de cette section, nous proposons une étude bibliographique sur les stratégies d'adaptation du gain pour les observateurs à grand gain de type de Luenberger (ce sont des observateurs dont la matrice de gain est calculée hors ligne, la plupart du temps par une méthode de placement de pôle). Dans le cas des observateurs de type de Kalman, la méthode de modification du gain permet de changer la structure des matrices  $Q$  et  $R$ . C'est pour cette raison que l'adaptation du gain peut être vue comme une adaptation des matrices de covariance de l'état et de bruit de mesure. Ces méthodes d'adaptation ont été largement abordées dans la littérature. D'autres méthodes qui reposent sur des règles empiriques sont proposées dans la littérature à savoir :

- Méthodes statistiques [Yu, 2005]
- Observateurs basés sur des algorithmes génétiques [La Moine et al., 1995]
- Observateurs basés sur les réseaux de neurones [Stubberud et al., 1998]
- Approches basées sur la logique floue [Jwo et Chang, 2007].

Bien que l'efficacité de ces méthodes soit avérée, peu ou aucune preuve de convergence n'a été mentionnée.

#### a. Adaptation du gain pour les observateurs de type Luenberger

Dans un article publié en 1997, Bullinger et Allogower ont proposé un observateur à grand gain de Luenberger ayant un gain variable [Bullinger et Allogower, 1997]. La structure de cet observateur est inspirée d'un travail précédent de [Tornambè, 1992]. La procédure d'adaptation définie dans cette étude se base sur une suite strictement croissante des éléments qui ne permettent pas au grand gain  $\theta$  de décroître. Ainsi, cet observateur est plus adapté à la résolution d'un problème de réglage des paramètres plutôt qu'un problème de réduction de bruit.

Les principales différences entre cet observateur et l'observateur adaptatif proposé par [Boizot et al, 2010], et appliqué pour l'estimation d'état d'un modèle ASM1 réduit dans le chapitre suivant, résident tout d'abord dans le fait que l'observateur utilisé dans notre travail est un observateur basé sur un filtre de Kalman, ce qui implique la prise en compte de l'évolution de la matrice de Ricatti. La seconde distinction est qu'ici, nous abordons le problème de la réduction du bruit lorsque l'estimation est suffisamment bonne.

Le deuxième observateur non-linéaire de type de Luenberger à gain adaptatif est proposé par Praly et ses collaborateurs. Une description détaillée des méthodes de construction de cet observateur peut être trouvée dans l'article [Andrieu et al., 2009]. Le travail proposé utilise une forme particulière de la forme canonique d'observabilité représentée par l'équation (2.23). L'observateur  $a$ , à peu près, la même structure qu'un observateur grand gain de Luenberger. Une légère différence réside dans le fait que le gain d'adaptation est une fonction de l'erreur de sortie. La stratégie adoptée dans leur travail ne s'appuie pas sur un champ de vecteur globalement Lipchitz, ce qui implique un calcul hors ligne du gain  $\theta$  à l'aide d'une certaine fonction d'adaptation. Cependant cette dernière n'est pas une fonction facile à trouver.

L'évolution du gain de l'observateur ne dépend pas de la convergence de l'erreur d'estimation. Par conséquent, l'observateur peut converger et le gain rester encore élevé en produisant une amplification de bruit. Le système utilisé dans le chapitre suivant présente l'avantage qu'il soit globalement Lipchitz. En outre, lorsque l'état estimé converge vers l'état réel, le gain diminue vers 1 (ou à la plus faible valeur fixée par l'utilisateur). Ainsi un gain élevé n'est plus nécessaire.

Le troisième exemple d'observateur de type de Luenberger est élaboré par [Ahrens et Khalil, 2009]. Le gain de cet observateur commute entre deux valeurs:

- la première valeur permet une reconstruction rapide de l'état,
- une deuxième rend l'observateur plus efficace au filtrage de bruit.

Cependant le schéma de commutation entre ces deux valeurs présente deux restrictions principales :

- La valeur de  $\theta$  devrait changer chaque fois qu'une erreur d'estimation trop importante est détectée,
- La valeur de  $\theta$  ne devrait pas changer dans le cas d'un dépassement. Lors d'un dépassement, la situation suivante peut survenir : la trajectoire estimée croise la trajectoire réelle, mais le système n'a pas encore convergé. Le passage d'un gain élevé à un gain faible dans ce cas n'est pas souhaitable.

L'observateur de Luenberger n'a pas les mêmes propriétés locales qu'un filtre de Kalman étendu à savoir : les bonnes propriétés de filtrage de bruit et la convergence pour des erreurs initiales faibles. C'est pour cette raison qu'on opte pour un observateur de Kalman étendu grand gain ayant un gain variable  $\theta$  pour être plus efficace en face du bruit.

### *b. Filtre de Kalman adaptatif*

Dans ce paragraphe, nous nous intéressons au problème d'adaptation du grand-gain  $\theta$  d'un filtre de Kalman. Rappelons que le paramètre de grand gain  $\theta$  sert à modifier les matrices de covariance  $Q$  et  $R$ . Il n'existe aucune preuve de convergence du filtre de Kalman étendu lorsque l'état estimé est éloigné initialement.

Les stratégies d'adaptation des matrices et du filtre de Kalman ont fait l'objet de nombreux travaux à la fois dans le cas linéaires et non-linéaires.

Pour les systèmes linéaires les lecteurs sont invités à lire les références suivantes [Gelb, 1974 ; Maybeck, 1982]. Dans le cas non-linéaire, les majorités des stratégies d'adaptation sont proposées pour les systèmes à temps discrets.

La première stratégie que nous présentons a été proposée par Pappas et Doss [Pappas et Doss, 1988]. Bien qu'ils ne considèrent pas le problème d'observabilité, il est néanmoins une préoccupation sous-jacente de leur travail :

- Lorsque le système n'est plus observable, un lent observateur est nécessaire pour donner une estimation précise et pour réduire le bruit.
- Lorsque le système devient observable (suite à un changement de point de fonctionnement ou un changement des caractéristiques physiques de processus), un observateur rapide peut être considéré pour assurer la convergence rapide des états estimés vers ceux réels.

La mise en œuvre de cette approche d'adaptation consiste en deux observateurs en parallèle. Les changements d'états sont considérés comme des défauts.

La deuxième stratégie consiste à concevoir un exemple d'observateur non-linéaire avec des valeurs différentes de  $Q$  et  $R$ . Ils sont utilisés en parallèle et la valeur estimée finale est choisie parmi toutes les estimations disponibles. La minimisation de l'innovation est l'un des plus simples critères de choix présenté par les auteurs [Bradshaw et al., 1997].

La troisième approche que nous citons est proposée dans [Boutayeb et Aubry, 1999]. Cette procédure d'adaptation est différente des stratégies données ci-dessus, dans le sens que les stratégies considérées n'ont pas les mêmes propriétés d'observabilité. En effet, dans le cas de leur travail, Boutayeb et Aubry, considèrent des systèmes  $N$ -localement uniformément observables. Cela implique que la classe des systèmes pour lequel l'observateur peut converger est plus grand que celui considéré dans le présent travail. Cet observateur est applicable même au processus qui ne peut pas être mis dans les formes canoniques d'observabilité définies précédemment (section 2.3.4). Cependant l'inconvénient de cette méthode d'adaptation est que l'observateur converge localement et asymptotiquement (l'erreur d'estimation n'est plus majorée par une exponentielle).

La stratégie d'adaptation de l'observateur adaptatif sujet de chapitre suivant est très similaire à celle là.

### c. Observateur proposé par Busvelle et Gauthier

Dans l'article [Busvelle et Gauthier, 2002], les auteurs proposent un observateur adaptatif de Kalman étendu grand gain. Cet observateur possède un gain élevé à  $t = 0$ , et peut diminuer à 1 par la suite. En d'autres termes l'observateur évolue d'un observateur grand gain qui permet une convergence rapide de la trajectoire estimée à vers un filtre de Kalman étendu qui conduit à un lissage du bruit. Cet article est le point de départ de l'observateur proposé par Boizot et al [Boizot et al., 2010] et appliqué pour l'estimation d'état d'un modèle ASM1 réduit au chapitre 3.

#### Définition 2.12

$$\begin{cases} \dot{z}(t) = A(u)z + b(z, u) - S^{-1}C'R^{-1}(Cz - y) \\ \dot{S}(t) = -(A(u) + b^*(z, u))'S - S(A(u) + b^*(z, u)) + C'R^{-1}C - SQ_\theta S \\ \dot{\theta}(t) = \lambda(1 - \theta) \end{cases} \quad (2.32)$$

Où  $Q_\theta = \theta^2 \Delta^{-1}Q\Delta^{-1}$ ,  $\Delta = \text{diag}(\{1, \frac{1}{\theta}, \dots, \frac{1}{\theta^{n-1}}\})$ ,  $Q$  et  $R$  sont les deux matrices classiques (covariances des bruits) d'un filtre de Kalman étendu définis précédemment (section 2.3.6).  $\lambda$  est un paramètre positif fixé, alors que  $\theta$  est le paramètre habituel de l'observateur grand-gain, à ceci près qu'il est ici dépendant du temps et plus précisément décroissant exponentiellement à partir de  $\theta(0) > 1$  jusqu'à 1 et avec une vitesse dépendant de  $\lambda$ .

- Si  $\theta(0) = 1$ , alors  $\theta(t) \equiv 1$ , nous retrouvons le filtre de kalman étendu classique mis sous forme canonique d'observabilité. Ainsi, selon les conditions initiales du système, l'observateur peut ne pas converger.
- Si  $\lambda = 0$ ,  $\theta(0) = \theta_0$  soit assez large et  $\lambda$  assez faible, alors  $\theta(t) = \theta$  garde une valeur élevée et nous retrouvons ainsi le filtre de Kalman étendu grand gain tel que représenté par l'équation (2.31).

L'idée générale de cet observateur est de choisir au début une valeur  $\theta(0)$  assez élevée et de définir une valeur de  $\lambda$  assez faible de telle sorte que l'observateur puisse converger rapidement. Lorsque l'état estimé est proche de l'état réel,  $\theta$  devient trop faible et la propriété de convergence locale d'un observateur de Kalman étendu peut garantir que l'erreur d'estimation restera petite.

#### Théorème 2.9

Il existe  $\lambda_0 > 0$  tel que pour tout  $0 \leq \lambda \leq \lambda_0$ , il existe  $\theta_0$  dépendant de  $\lambda$  tel que pour tout  $\theta(0) > \theta_0$ , pour tout  $S(0) \geq c \text{Id}$ , pour tout compact  $K \subset \mathbb{R}^n$ , pour tout  $z(0) \in K$  alors si on pose  $\varepsilon(t) = z(t) - x(t)$ , on a pour tout  $t \geq 0$

$$\|\varepsilon(t)\|^2 \leq R(\lambda, c)e^{-at} \Lambda(\theta(0), t, \lambda) \|\varepsilon(0)\|^2$$

$$\text{Où } \Lambda(\theta(0), t, \lambda) = \theta(0)^{2(n-1) + \frac{\alpha}{\lambda}} e^{-\frac{\alpha}{\lambda}\theta(0)(1-e^{-\lambda t})}$$

et où  $a$  est un réel positif alors que  $R(\lambda, c)$  est une fonction décroissante de  $c$ .

Ce théorème montre que l'observateur converge exponentiellement quelque soit les erreurs initiales. Cet observateur peut être utilisé pour estimer la trajectoire d'un système dont on ne connaît pas l'état initial mais en connaissant parfaitement les équations d'évolutions. Néanmoins, il

est clair qu'il est non persistant puisque après un certain temps on retrouve les équations d'un filtre de Kalman étendu lorsque  $\theta(t)$  se rapproche de 1. Afin d'obtenir un observateur persistant et capable d'estimer l'état du système même en cas de grandes perturbations (non modélisées), les auteurs proposent d'utiliser plusieurs observateurs de type (2.32) en faisant de sorte qu'à chaque instant cohabitent au moins un observateur grand gain et un observateur de Kalman étendu classique dont chacun d'eux est initialisé à différents moments.

Le principe est de considérer une famille d'observateurs  $= 1, \dots, N$ , tous de la même forme mais tels que à chaque instant de la forme  $(i + \frac{j-1}{N})T, j = 1, \dots, N, i \in N$ , le  $j^{\text{ème}}$  observateur soit réinitialisé en  $(z(0), S(0), \theta(0))$ . Ainsi, chaque observateur a « une durée de vie » de  $T$  et à tout instant, il existe un observateur qui a été réinitialisé depuis un intervalle de temps n'excédant pas  $\frac{T}{N}$ . En choisissant  $T$  et  $N$ , il aura à chaque instant :

- Au moins un observateur de type de Kalman étendu classique
- Au moins un observateur de type Kalman étendu grand gain

Ainsi l'observateur qui prédit le mieux le comportement du système réel, qui minimise donc l'erreur d'estimation, est choisi. Comme illustré dans l'article [Busvelle et Gauthier, 2002], l'application de cet observateur à un modèle de colonne à distiller à 5 plateaux théoriques montrent que :

- Le choix des paramètres est plus simple que pour le filtre de Kalman étendu : en effet, ce dernier nécessite de choisir  $Q$  et  $R$  de façon à trouver un compromis entre un comportement efficace en l'absence ou non de perturbations. Ici, l'introduction de nouveaux paramètres permet de choisir  $Q$  et  $R$  de façon à avoir de bonnes performances en l'absence de perturbations, les autres paramètres garantissant la bonne performance lors des perturbations.
- Les observateurs sont indépendants les uns des autres,
- La multiplicité des observateurs permet de fournir des informations sur les instants des apparences des perturbations, ce qui forme un outil performant de diagnostic.

Malgré l'efficacité de cette procédure, elle présente les inconvénients suivants :

- La construction est dépendante de temps,
- Elle nécessite un temps de calcul élevé puisque on a besoin de  $N$  observateurs en parallèle.
- Le paramètre  $\lambda$  est choisi assez faible de sorte qu'après une perturbation, on ne pourra pas revenir aussi vite à un filtre de Kalman étendu classique,
- Le critère de choix, pour sélectionner la meilleure prédiction du comportement du système, n'est pas justifié théoriquement,

## 2.4. Conclusion

Ce deuxième chapitre a été consacré essentiellement à quelques rappels indispensables à toute stratégie de diagnostic et d'observation et nécessaires à la compréhension de ce mémoire.

Une grande partie de ce chapitre a été réservée à un rappel des notions d'observabilité, élaborées par Gauthier et Kupka, qui consistent à une mise du système sous certaine forme canonique d'observabilité. Ainsi, nous pouvons appliquer tous les types d'observateurs grand gain : ceux de type de Luenberger et ceux de type de Kalman. Les premiers sont des observateurs globaux à convergence exponentielle et ne sont appliqués qu'après une étude d'observabilité. De l'autre côté, le filtre de Kalman étendu est l'observateur le plus largement utilisé depuis longtemps, même pour des procédés complexes. Afin de réunir les avantages de ces deux types d'observateurs, un filtre de Kalman étendu grand gain a été proposé par Boizot et al. [Boizot et al., 2010], pour un système non-linéaire MISO (Multi-entrées Mono-sortie). En effet, cet observateur adaptatif unit les avantages d'un filtre de Kalman étendu (FKE), en présence de bruit et un observateur de Kalman étendu grand gain (OKE grand gain) dans le cas d'une large variation.

Une présentation de cet observateur ainsi que son application à une station d'épuration biologique à boues activées fera l'objet du chapitre suivant.