



**HAL**  
open science

# Diagnostic des défauts des machines asynchrones par reconnaissance des formes

Roland Casimir

► **To cite this version:**

Roland Casimir. Diagnostic des défauts des machines asynchrones par reconnaissance des formes. Autre. Ecole Centrale de Lyon, 2003. Français. NNT : . tel-00139706

**HAL Id: tel-00139706**

**<https://theses.hal.science/tel-00139706>**

Submitted on 3 Apr 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THESE

Présentée devant

**L'ÉCOLE CENTRALE DE LYON**

pour obtenir le grade de

**DOCTEUR**

(arrêté du 30/03/1992)

**Spécialité: Génie Électrique**

Préparée au sein de

L'ÉCOLE DOCTORALE

ÉLECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE

DE LYON

Par

**Roland CASIMIR**

---

**DIAGNOSTIC DES DEFAUTS DES MACHINES ASYNCHRONES  
PAR RECONNAISSANCE DES FORMES**

---

*Soutenue le 18 décembre 2003 devant la commission d'examen*

Jury : MM.

B. DUBUISSON	Professeur des Universités HEUDIASYC - Compiègne	Rapporteur
J. C. HAPIOT	Professeur des Universités LEEI - Toulouse	Rapporteur
G. CLERC	Professeur des Universités CEGELY - Lyon	Examinateur
G. BARAKAT	Maître de Conférence (HDR) GREAH – Le Havre PSA Peugeot Citroën	Examinateur
A. BOATAS	Responsable thème Traitement du Signal et Diagnostic	Examinateur
E. BOUTLEUX	Maître de Conférence CEGELY - Lyon	Examinateur

*A mes chers parents Cordélia et Augustin*

*A toute ma famille,*

*mes amis,*

*A Judith.*

## **REMERCIEMENTS**

*Les travaux de recherche développés dans ce mémoire ont été réalisés au sein du Centre de Génie Electrique de Lyon (CEGELY – UMR 5005 CNRS) sur le site de l'Ecole Centrale de Lyon.*

*Je remercie très sincèrement Messieurs les Professeurs Alain NICOLAS, directeur du CEGELY, et Philippe AURIOL, directeur du département EEA, de m'avoir accueilli dans ce laboratoire.*

*J'exprime ma profonde gratitude et ma reconnaissance envers Messieurs Guy CLERC, Professeur à l'Université Claude Bernard Lyon 1, et Emmanuel BOUTLEUX, maître de Conférences à l'Ecole Centrale de Lyon, pour l'aide précieuse qu'ils m'ont apporté dans l'élaboration de ces travaux. Je leur remercie pour leurs conseils, leur soutien, leur disponibilité et surtout leurs grandes qualités humaines.*

*Je remercie vivement Messieurs Bernard DUBUISSON, Professeur à l'Université de Technologie de Compiègne, et Jean – Claude HAPIOT, Professeur à l'Institut National Polytechnique de Toulouse, de m'avoir fait l'honneur d'accepter d'être les rapporteurs de cette thèse.*

*Je tiens également à adresser mes remerciements à Monsieur Georges BARAKAT, maître de Conférences HDR à l'Université du Havre, et à Monsieur Armand BOATAS du groupe PSA Peugeot Citroën, de m'avoir fait l'honneur participer au jury.*

*Durant ces années de doctorat, les membres administratifs et techniques du département et du laboratoire ont toujours su répondre présent à mes nombreuses sollicitations. J'ai particulièrement apprécié leur disponibilité, leur sympathie, et leurs compétences. Qu'ils reçoivent toute ma gratitude !*

*Je ne saurai enfin oublier les jeunes docteurs et futur docteurs du laboratoire qui, durant ces années, ont entretenu une très bonne ambiance et ont fait preuve de solidarité. Je souhaite aux doctorants du succès dans leurs recherches ainsi que dans leurs vies professionnelles.*

## **RESUME**

Les progrès réalisés en matière d'électronique de puissance, de circuits de commande et en automatique ont contribué à l'utilisation grandissante des machines asynchrones dans les systèmes d'entraînements électriques. Le recours aux machines asynchrones est surtout lié à leur robustesse, leur puissance massique et à leur coût de fabrication. Leur maintenance et leur surveillance permettent de rentabiliser les installations. Il est donc important de développer des outils de diagnostic pour détecter de manière précoce les défauts pouvant apparaître dans ces machines.

Notre approche est basée sur l'utilisation des méthodes de reconnaissance des formes. Un vecteur de paramètres, appelé vecteur forme, est extrait de chacune des mesures effectuées sur la machine. Les règles de décisions utilisées permettent de classer les observations, décrites par le vecteur forme, par rapport aux différents modes de fonctionnement connus avec ou sans défaut.

Des défauts ont été créés au rotor et au stator de la machine asynchrone, alimentée soit à partir du réseau, soit par le biais d'un onduleur de tension.

Les procédures de décisions, basées sur la règle des k - plus proches voisins et sur le calcul direct des frontières, ont été utilisées pour détecter les défauts. Ces algorithmes ont montré l'efficacité de l'application de la reconnaissance des formes au diagnostic.

## **ABSTRACT**

Advances in power electronics, control circuits and automatic have contributed to an increasing use of induction motors in electrical drive systems. The large – scale utilization of induction motors is mainly due to their robustness, their power – weight ratio, and to their manufacturing cost. Therefore, it is important to develop diagnosis tools in order to detect earlier the faults, which can appear in these machines.

Our approach is based on pattern recognition methods. A vector of parameters, called pattern vector, is obtained from the measurements made on the machine. The decision rules enable to classify the observations described by the pattern vector. These classifications are made according to the different operating conditions, with or without fault.

Faults have been created on both the rotor and the stator sides of the induction machine. This one was fed either from the mains, or from a three – phase voltage inverter.

Fault detection has been made with decision procedures based on the k – nearest neighbors rule and on boundaries direct calculation. The results obtained with these algorithms have proved the efficiency of pattern recognition methods for diagnosis.

# SOMMAIRE

INTRODUCTION GENERALE.....	1
----------------------------	---

---

## CHAPITRE I : SYNTHÈSE DES MÉTHODES DE DIAGNOSTIC DE LA MACHINE ASYNCHRONE

---

<b>1.1 - Introduction .....</b>	<b>5</b>
<b>1.2 - Méthodes de diagnostic avec connaissance a priori.....</b>	<b>6</b>
1.2.1 Principe.....	6
1.2.2 Modèle de la machine asynchrone .....	7
1.2.2.1 Transformation triphasée / diphasée.....	7
1.2.2.2 Transformation de Park .....	8
1.2.2.3 Choix du référentiel.....	9
1.2.3 Application au diagnostic.....	10
1.2.3.1 Suivi du couple électromagnétique.....	11
1.2.3.2 Identification des paramètres de la machine asynchrone .....	13
1.2.3.2.1 Équations d'état.....	13
1.2.3.2.2 Modèle pour l'identification.....	15
1.2.3.3 Paramètres représentatifs des défauts .....	15
1.2.4 Conclusion.....	19
<b>1.3 - Méthodes de diagnostic sans connaissance a priori .....</b>	<b>22</b>
1.3.1 Principe.....	22
1.3.2 Suivi des grandeurs mesurables .....	22
1.3.2.1 Analyse des courants .....	22
1.3.2.1.1 Excentricité .....	22
1.3.2.1.2 Défaillance des paliers .....	23

1.3.2.1.3	Rupture de barres ou de portions d'anneaux.....	25
1.3.2.2	Analyse des flux .....	25
1.3.2.2.1	Court-circuit entre spires au stator.....	26
1.3.2.2.2	Cas d'une dissymétrie au rotor.....	28
1.3.2.3	Analyse du couple électromagnétique.....	28
1.3.2.4	Analyses vibratoires de la machine .....	29
<b>1.4 -</b>	<b><i>Synthèse générale des résultats</i></b> .....	<b>31</b>
1.4.1	Champ d'application des méthodes de diagnostic .....	31
1.4.2	Conclusion.....	34

---

## **CHAPITRE II :**

## **RECONNAISSANCE DES FORMES ET DIAGNOSTIC**

---

<b>2.1 -</b>	<b><i>Introduction</i></b> .....	<b>36</b>
<b>2.2 -</b>	<b><i>Analyse des observations</i></b> .....	<b>39</b>
2.2.1	Détermination de l'espace de représentation .....	39
2.2.2	Réduction de l'espace de représentation.....	39
2.2.2.1	Analyse en Composantes Principales (ACP).....	40
2.2.2.2	Sélection de paramètres .....	45
<b>2.3 -</b>	<b><i>Détermination de l'espace de décision</i></b> .....	<b>50</b>
2.3.1	Méthodes de classification hiérarchiques.....	51
2.3.2	Méthodes de classification par partition.....	52
<b>2.4 -</b>	<b><i>Choix d'une procédure de détection</i></b> .....	<b>56</b>
2.4.1	Méthodes paramétriques.....	58
2.4.1.1	Règle de Bayes .....	58
2.4.1.2	Rejet d'ambiguïté .....	58
2.4.1.3	Rejet de distance .....	59
2.4.2	Méthodes non paramétriques.....	61
2.4.2.1	Principe des méthodes non paramétriques.....	61



2.4.2.2	Règle des $k$ – plus proches voisins ( $k$ - ppv).....	62
2.4.2.3	Rejet d'ambiguïté .....	64
2.4.2.4	Rejet de distance .....	65
2.4.3	Calcul direct des frontières.....	65
2.4.3.1	Présentation du cas linéaire .....	66
2.4.3.2	Discriminateur linéaire de Fisher.....	69
2.4.3.3	Décision entre plusieurs classes.....	71
<b>2.5 -</b>	<b><i>Evaluation des performances de la procédure</i></b> .....	<b>73</b>
2.5.1.	Critère de qualité.....	73
2.5.2.	Méthodes de test .....	74
<b>2.6 -</b>	<b><i>Exploitation du système de diagnostic</i></b> .....	<b>75</b>
<b>2.7 -</b>	<b><i>Conclusion</i></b> .....	<b>75</b>

---

**CHAPITRE III :**                      APPLICATION AU DIAGNOSTIC DE LA MACHINE ASYNCHRONE

---

<b>3.1 -</b>	<b><i>Introduction</i></b> .....	<b>78</b>
<b>3.2 -</b>	<b><i>Description du procédé</i></b> .....	<b>78</b>
3.2.1	Présentation du banc moteur .....	80
3.2.2	Moyens de mesures .....	81
3.2.3	Différents modes de fonctionnement .....	83
<b>3.3 -</b>	<b><i>Recherche de signatures</i></b> .....	<b>84</b>
3.3.1	Prétraitement des signaux et calcul des paramètres .....	86
3.3.1.1	Données temporelles.....	86
3.3.1.1.1	Transformation triphasée – diphasée (axes $\alpha$ , $\beta$ liés au stator).....	87
3.3.1.1.2	Puissances .....	88
3.3.1.1.3	Calcul sur $I_{s\alpha}$ et $I_{s\beta}$ .....	91
3.3.1.2	Données fréquentielles.....	95

3.3.1.2.1	Analyse du courant stator .....	95
3.3.1.2.2	Amplitude des raies aux fréquences caractéristiques.....	96
3.3.1.2.3	Calcul énergétique .....	99
3.3.1.3	Composantes directes et inverses .....	100
3.3.1.3.1	Impédance directe.....	100
3.3.1.3.2	Impédance inverse .....	101
3.3.2	Choix du vecteur forme .....	102
3.3.2.1	Sélection par le critère de Fisher .....	104
3.3.2.2	Sélection par la méthode SBS .....	106
3.3.3	Visualisation par analyse en composantes principales.....	108
3.3.4	Conclusion.....	111
<b>3.4 -</b>	<b><i>Procédure de décision</i></b> .....	<b>112</b>
3.4.1	Choix du vecteur forme .....	113
3.4.2	Décision par la règle des k – ppv .....	116
3.4.2.1	Phase d'exploitation : classification sans rejet.....	121
3.4.2.2	Phase d'exploitation : classification avec rejet .....	123
3.4.3	Décision à partir des frontières linéaires .....	127
3.4.3.1	Résultats de classification.....	129
3.4.3.2	Extension à la notion de rejet .....	132
3.4.4	Comparaison des procédures de décision.....	135
3.4.5	Cas de l'alimentation avec onduleur.....	135
3.4.6	Conclusion.....	139
	<b>CONCLUSION GENERALE</b> .....	<b>141</b>
	<b>BIBLIOGRAPHIE</b> .....	<b>146</b>

# NOTATIONS

## Notations liées à la machine asynchrone

---

[C]	Matrice de Concordia (repère diphasé de Park fixe lié au stator)
$C_{em}$	Couple électromagnétique
$C_r$	Couple résistant imposé par la charge du moteur
$C_u$	Couple utile en sortie du moteur
$f_s$	Fréquence des courants statoriques $f_s = \frac{\omega_s}{2\pi}$
$f_t$	Coefficient de frottement
$f_r$	Fréquence de rotation du rotor $f_r = \frac{\omega}{2\pi} = \frac{(1-g)f_s}{p}$
$g$	Glissement du moteur $g = \frac{p\omega_r}{\omega_s}$
$i_{s\alpha,\beta}, i_{r\alpha,\beta}$	Courant statorique (respectivement rotorique) suivant l'axe $\alpha$ (ou $\beta$ )
$i_{sd,q}, i_{rd,q}$	Courant statorique (respectivement rotorique) suivant l'axe d (ou q)
$J$	Moment d'inertie
$L_{fs}$	Inductance de fuite au stator
$l_{ms}$	Inductance magnétisante
$L_s$	Inductance propre cyclique du stator
$L_r$	Inductance propre cyclique du rotor
$L_m$	Inductance mutuelle entre le stator et le rotor
$N_r$	Nombre d'encoches (ou de barres) au rotor
$\omega_s$	Pulsation des courants statoriques
$\omega$	Vitesse angulaire du rotor

$\omega_r$	Pulsation des courants rotoriques $\omega_r = \frac{\omega_s}{p} - \omega$
$p$	Nombre de paires de pôles
$\varphi_{sd,q}, \varphi_{rd,q}$	Flux statorique (respectivement rotorique) suivant l'axe d (ou q)
$\varphi_{sa,\beta}, \varphi_{ra,\beta}$	Flux statorique (respectivement rotorique) suivant l'axe $\alpha$ (ou $\beta$ )
$R_s, R_r$	Résistance par phase statorique, respectivement rotorique
$\sigma$	Coefficient de fuite total ( $\sigma = 1 - \frac{L_m^2}{L_s L_r}$ )
$\theta_s$	Angle électrique entre le stator et l'axe d
$\theta_r$	Angle électrique entre le rotor et l'axe d
$T_s, T_r$	Constante de temps électrique du stator, respectivement du rotor
[T]	Matrice de transformation de Park
$V_{sa,\beta}, V_{ra,\beta}$	Tension statorique (respectivement rotorique) suivant l'axe $\alpha$ (ou $\beta$ )
$V_{sd,q}, V_{rd,q}$	Tension statorique (respectivement rotorique) suivant l'axe d (ou q)
[X <sub>r</sub> ]	Notation vectorielle pour un système triphasé au rotor
[X <sub>s</sub> ]	Notation vectorielle pour un système triphasé au stator
$\underline{x}$	Notation complexe
$X_{ra,b,c}$	Composantes rotoriques de la grandeur X <sub>r</sub> de la phase a (respectivement b, c)
$X_{sa,b,c}$	Composantes statoriques de la grandeur X <sub>s</sub> de la phase a (respectivement b, c)

### *Notations liées à la Reconnaissance des Formes*

---

<b>A</b>	Notation matricielle
<b>det(A),  A </b>	Déterminant de la matrice <b>A</b>
<b>A<sup>-1</sup></b>	Matrice inverse de la matrice <b>A</b>
<b>A<sup>t</sup></b>	Matrice transposée de la matrice <b>A</b>
<b>trace(A)</b>	Trace de la matrice <b>A</b>

$\alpha$	Un paramètre extrait de l'analyse des signaux. Si $\alpha$ est pertinent alors $\alpha \leftrightarrow x_j$
$C_d^{d'}$	Combinaison de $d'$ éléments parmi $d$
$d(\mathbf{x}, \mathbf{y})$	Calcul de la distance entre deux vecteurs $\mathbf{x}$ et $\mathbf{y}$
$d_E(\mathbf{x}, \mathbf{y})$	Distance Euclidienne
$d_M(\mathbf{x}, \mathbf{y})$	Distance de Mahalanobis
$d$	Dimension de l'espace
$f(\mathbf{x} \omega_i)$	Densité de probabilité de $\mathbf{x}$ dans $\omega_i$
$f(\mathbf{x})$	Densité de mélange
$\Gamma$	Pourcentage d'inertie ou qualité de la représentation dans le sous – espace de dimension $d'$
$g_{ij}(\mathbf{x})$	Décision linéaire pour $\mathbf{x}$ entre $\omega_i$ et $\omega_j$
$J, J', J''$	Critères
$k$	Nombre de plus proches voisins d'un vecteur $\mathbf{x}$ (règle des $k - \text{ppv}$ )
$k_i$	Nombre de plus proches voisins de $\mathbf{x}$ (parmi $k$ ) dans la classe $\omega_i$
$\lambda_j$	$j^{\text{ième}}$ valeur propre de $\mathbf{S}$ ( $j = 1, \dots, d$ )
$\mathbf{m}$	Centre de gravité de l'ensemble
$m_j$	$j^{\text{ième}}$ composante de $\mathbf{m}$
$\mathbf{m}_i$	Centre de gravité de la classe $\omega_i$
$M$	Nombre de classes
$N$	Nombre de vecteurs dans l'ensemble d'apprentissage
$N_i$	Nombre d'échantillons dans la classe $\omega_i$
$\nabla$	Opérateur gradient
$\omega_0$	Classe fictive des vecteurs (ou individus) rejetés en ambiguïté
$\omega_d$	Classe fictive des vecteurs rejetés en distance
$\omega_i$	Dénomination de la classe $i$ (avec $i = 1, \dots, M$ )
$\Omega_i$	Région de l'espace où est située la classe $\omega_i$
$\text{Pr}(\omega_j   \mathbf{x})$	Probabilité a posteriori pour qu'une observation $\mathbf{x}$ provienne de la classe $\omega_j$

$\Pr(\omega_i)$	Probabilité a priori de la classe $\omega_i$
$\sigma_j^2$	Composante $j$ de la variance des vecteurs
$\sum_{i=1}^M ()$	Somme sur l'indice $i$
$\mathbf{S}_w$	Matrice de variance intra – classe de l'ensemble
$\mathbf{S}_B$	Matrice de variance inter – classes de l'ensemble
$\mathbf{S}$	Matrice de variance – covariance de l'ensemble
$\mathbf{S}_i$	Matrice de variance – covariance de la classe $\omega_i$
$T_d$	Seuil de rejet de distance
$\ \mathbf{v}\ $	Norme de vecteur
$\mathbf{v}_j$	$j^{\text{ième}}$ vecteur propre de $\mathbf{S}$ ( $j = 1, \dots, d$ )
$\mathbf{w}, w_0$	Respectivement vecteur poids et seuil définissant la frontière linéaire séparant deux classes données $\omega_i$ et $\omega_j$
$\mathbf{x}$	Vecteur de dimension ( $d \times 1$ )
$\mathbf{x}'$	Voisin de $\mathbf{x}$
$x_j$	$j^{\text{ième}}$ composante du vecteur forme ( $j = 1, \dots, d$ )
$\mathbf{x}^p$	Projection de $\mathbf{x}$
$\mathbf{x}_k$	$k^{\text{ième}}$ vecteur (ou échantillon ou encore observation) de l'ensemble d'apprentissage
$x_{kj}$	$j^{\text{ième}}$ composante de $\mathbf{x}_k$
$\mathbf{X}$	Ensemble (ou matrice) des $N$ vecteurs de l'ensemble ; matrice des données
$\mathbf{x}_{ij}$	$j^{\text{ième}}$ vecteur de la classe $\omega_i$ ( $j = 1, \dots, N_i$ et $i = 1, \dots, M$ )
$\mathbf{X}^p$	Projection de la matrice des données $\mathbf{X}$ sur les composantes principales

# INDEX DES FIGURES

## CHAPITRE I

- Fig. 1. 1.** *Principe du diagnostic par modèle.*
- Fig. 1. 2.** *Photo d'une machine asynchrone 5,5 kW avec le stator et le rotor qui la constituent.*
- Fig. 1. 3.** *Représentation schématique des enroulements de la machine dans le plan  $a, b, c$ .*
- Fig. 1. 4.** *Schéma de principe de la méthode du modèle de référence.*
- Fig. 1. 5.** *Représentation schématique des encoches stator [Schaeffer, 99 a]. Localisation des spires en court – circuit. La bobine fictive  $B_{cc}$  est représentée dans le repère  $\alpha, \beta$*
- Fig. 1. 6.** *Représentation de l'excentricité statique et dynamique.*
- Fig. 1. 7.** *Vue en coupe de la partie supérieure d'un roulement à bille [Benbouzid, 99] [Schoen, 95 a] [Schoen, 95 b]*
- Fig. 1. 8.** *Mesures vibratoires sur la machine asynchrone.*

## CHAPITRE II

- Fig. 2.1.** *Caractérisation d'une observation par un vecteur forme représenté par un point dans  $\mathbb{R}^d$  (avec  $d = 3$ ).*
- Fig. 2.2.** *Objectif en Reconnaissance des Formes : associer une nouvelle observation  $\mathbf{x}$  à l'une des classes.*
- Fig. 2.3.** *Phase d'analyse : caractérisation de l'état du système (à un instant donné) par le vecteur forme.*
- Fig. 2.4.** *Illustration de l'information apportée par les composantes principales. [<http://www.ulg.ac.be/pedaexpe/cours/glosaire/acp.htm>]*
- Fig. 2.5.** *Exemple d'analyse permettant de choisir la dimension du vecteur forme.*
- Fig. 2.6.** *Schéma de principe de la sélection de paramètres par la méthode SBS, avec  $d=5$  et  $d'=2$ .*
- Fig. 2.7.** *Schéma de principe de la classification hiérarchique ascendante. L'ensemble d'apprentissage est constitué de 9 observations.*
- Fig. 2.8.** *Résultat de la partition aléatoire de l'ensemble. 46 échantillons ont été affectés à la classe  $\omega_1$  ("0") et 54 à la classe  $\omega_2$  ("+").*

- Fig. 2.9.** *Résultat de la coalescence HCM. Après 7 itérations : partition stricte en deux classes  $\omega_1$  ("0") et  $\omega_2$  ("+") de 50 échantillons et centres de gravité obtenus(■).*
- Fig. 2.10.** *Variation de la trace de la matrice de dispersion intra-classe au cours de la coalescence.*
- Fig. 2.11.** *Observation située près de la frontière de décision entre deux classes.*
- Fig. 2.12.** *Application du rejet de distance. La nouvelle observation  $\mathbf{x}$  ne correspond à aucune des classes connues.*
- Fig. 2.13.** *Illustration de l'affectation par la règle des  $k$  – ppv, avec  $k = 7$ .*
- Fig. 2.14.** *Géométrie de la fonction discriminante linéaire dans un espace de dimension 2*
- Fig. 2.15.** *Matérialisation de la frontière de décision entre les deux classes par la projection des points – représentation en dimension 2.*
- Fig. 2.16.** *Matérialisation de la frontière de décision entre les deux classes par la projection des points – représentation en dimension 3.*
- Fig. 2.17.** *Généralisation de l'analyse discriminante au problème multiclassés.*

### CHAPITRE III

- Fig. 3. 1.** *Photo du banc moteur 5,5 kW dédié au diagnostic.*
- Fig. 3. 2.** *Schéma de principe de l'acquisition des signaux.*
- Fig. 3. 3.** *Photos des rotors utilisés.*
- Fig. 3. 4.** *Principe de sélection de la signature (vecteur forme) utilisé dans l'application.*
- Fig. 3. 5.** *Evolution des composantes  $I_{s\alpha}$  et  $I_{s\beta}$  des courants stators pour un rotor sain (a) et pour un rotor ayant 4 barres cassées (b).*
- Fig. 3. 6.** *Caractéristique normalisée  $Q'(P')$  obtenue pour la machine saine, pour une charge de 3 kW en alimentation réseau.*
- Fig. 3. 7.** *Variations de la dispersion  $\zeta$  sur l'ensemble des acquisitions à vide et en charge. Machine saine (\*). Rotor en défaut (o). Déséquilibre stator ( $\Delta$ ).*
- Fig. 3. 8.** *Variations des puissances moyennes. Machine saine (\*). Rotor en défaut (o). Déséquilibre stator ( $\Delta$ ).*
- Fig. 3. 9.** *Variations des écarts – types des courants  $I_{s\alpha}$  et  $I_{s\beta}$ . Machine saine (\*). Rotor en défaut (o). Déséquilibre stator ( $\Delta$ ).*



- Fig. 3. 10.** *Variations de l'écart – type de  $I_{s\alpha,\beta}$ . Machine saine (\*). Rotor en défaut (o). Déséquilibre stator ( $\Delta$ ).*
- Fig. 3. 11.** *Caractéristique  $I_{s\beta}(I_{s\alpha})$  normalisée pour le fonctionnement à vide.*
- Fig. 3. 12.** *Caractéristique  $I_{s\beta}(I_{s\alpha})$  normalisée pour le fonctionnement en charge.*
- Fig. 3. 13.** *Déformation de la caractéristique  $I_{s\beta}(I_{s\alpha})$  normalisée pour l'ensemble des acquisitions.*
- Fig. 3. 14.** *Spectre du courant  $I_{sd}$ , fonctionnement en charge. Rotor sain (a), rotor en défaut (b).*
- Fig. 3. 15.** *Variation des raies correspondant aux fréquences recherchées. Fonctionnement à vide.*
- Fig. 3. 16.** *Variation des raies correspondant aux fréquences recherchées. Fonctionnement en charge.*
- Fig. 3. 17.** *Variations de la puissance moyenne du courant. Machine saine (\*). Rotor en défaut (o). Déséquilibre stator ( $\Delta$ ).*
- Fig. 3. 18.** *Variation de l'impédance directe et de l'impédance inverse. Machine saine (\*). Rotor en défaut (+). Déséquilibre stator (x).*
- Fig. 3. 19.** *Variation du rapport de Fisher pour chaque paramètre.*
- Fig. 3. 20.** *Procédure de sélection des paramètres par la méthode SBS.*
- Fig. 3. 21.** *Valeur du critère  $J(E_k)$  à chaque étape de la sélection SBS.*
- Fig. 3. 22.** *Position des classes dans le sous – espace de dimension 3. Les paramètres sont sélectionnés avec le critère de Fisher.*
- Fig. 3. 23.** *Position des classes en dimension 3. Les paramètres sont sélectionnés avec l'algorithme SBS.*
- Fig. 3. 24.** *Ensemble d'apprentissage défini par le vecteur forme  $VF_1$ . Machine saine "■", trois barres cassées "●", et déséquilibre stator "◆".*
- Fig. 3. 25.** *Ensemble d'apprentissage défini par le vecteur forme  $VF_2$ . Machine saine "■", trois barres cassées "●", et déséquilibre stator "◆".*
- Fig. 3. 26.** *Illustration de la classification après calcul de la distance de Mahalanobis et de la distance Euclidienne.*
- Fig. 3. 27.** *Résultat de test sur les deux ensembles d'apprentissage. Les 60 échantillons de test sont sélectionnés aléatoirement.*

- Fig. 3. 28.** *Classification avant et après options de rejets des échantillons de test. Les vecteurs sont définis par  $VF_1$  (rapport de Fisher).*
- Fig. 3. 29.** *Classification avant et après options de rejets des échantillons de test. Les vecteurs sont définis par  $VF_2$  (algorithme SBS).*
- Fig. 3. 30.** *Analyse spectrale des mesures "Machine saine" et "3 barres cassées" à 100 % de charge.*
- Fig. 3. 31.** *Analyse spectrale des mesures "Machine saine" à 75 % et à 100 % de charge.*
- Fig. 3. 32.** *Schéma de principe de la procédure de diagnostic.*
- Fig. 3. 33.** *Résultats de diagnostic sans rejet, par les frontières linéaires, effectué sur les observations de test.*
- Fig. 3. 34.** *Récapitulatif du calcul effectué pour chaque observation vis à vis de sa position par rapport à l'ensemble des classes.*
- Fig. 3. 35.** *Résultats en terme de nombre d'échantillons mal classés après cent sélections aléatoires des individus de test. Décision par calcul direct des frontières.*
- Fig. 3. 36.** *Phase de décision sans rejet par les frontières linéaires pour les nouvelles observation. L'espace est caractérisée par  $VF_2$ .*
- Fig. 3. 37.** *Illustration de la discrimination linéaire avec rejet de distance.*
- Fig. 3. 38.** *Résultats de classification par les frontières linéaires après application du rejet de distance. Les observations de test sont définies par  $VF_2$ .*
- Fig. 3. 39.** *Calcul des distances entre les nouvelles observations et les hyperplans.*
- Fig. 3. 40.** *Tension générée par l'onduleur sur une phase de la machine, et courant de ligne correspondant.*
- Fig. 3. 41.** *Visualisation des données onduleur, caractérisées par les paramètres de  $VF_2$ , dans le meilleur sous – espace de dimension 3. 99,8 % de l'information est conservée.*
- Fig. 3. 42.** *Résultats de classification, par les  $k - ppv$  et le calcul des frontières, des individus supplémentaires issus des données onduleur.*
- Fig. 3. 43.** *Visualisation de l'ensemble des données onduleur dans l'espace factoriel. Le plan  $H$  donne la position éventuelle de la frontière séparant les classes 2 et 6.*



---

# Introduction générale

---

Les machines asynchrones, de par leur robustesse et leur rapport poids/puissance, sont largement utilisées en milieu industriel. Assurer leur continuité de fonctionnement nécessite la mise en place de programmes de maintenances préventive et corrective. En effet, la fiabilité et la sûreté de leur fonctionnement permettent en partie d'assurer la sécurité des personnes, la qualité du service et la rentabilité des installations.

Malheureusement, les contraintes nouvelles et l'intégration de ces machines dans des systèmes de conversion d'énergie de plus en plus complexes rendent le diagnostic plus difficile.

Il existe plusieurs procédures de diagnostic [Zwingelstein, 95]. Le choix d'une approche est lié à la connaissance que l'on souhaite acquérir sur le système, mais aussi à la complexité de ce système. Ainsi deux grandes familles de procédures sont utilisées dans le domaine du diagnostic en génie électrique : les méthodes de diagnostic à base de modèles analytiques et les méthodes sans modèle.

Les méthodes à base de modèles analytiques reposent sur le suivi des paramètres et des grandeurs de la machine, au moyen d'algorithmes d'observation. Elles détectent les défaillances en comparant l'évolution de l'écart entre le modèle et le processus réel. Le principal avantage de ces méthodes réside dans l'intégration d'une connaissance a priori du système et donc un filtrage de l'information.

Les méthodes sans modèle sont basées sur l'extraction d'informations par le biais du traitement des signaux mesurés. Les signaux mesurables (les courants, les tensions, la vitesse, les vibrations, la température, les émissions sonores) peuvent fournir des informations significatives sur les défauts. A partir de ces grandeurs caractéristiques du fonctionnement des machines électriques, la mise en œuvre de méthodes décisionnelles (par exemple à base de reconnaissance des formes) permet de concevoir des systèmes de surveillance ou des algorithmes performants au service du diagnostic. Les performances de ces méthodes sont étroitement liées à la pertinence des indicateurs de défauts retenus et à la finesse d'analyse des mesures.

---

Les travaux présentés dans ce mémoire portent sur le diagnostic d'une machine asynchrone en utilisant des méthodes de reconnaissance des formes (RDF) :

- Au niveau de l'application, l'étude est cadrée sur la détection de défauts électriques créés au rotor et au stator de la machine asynchrone. La machine est alimentée soit à partir du réseau, soit par le biais d'un onduleur de tension. Afin de ne pas « trop instrumenter » le processus à surveiller (ajout d'accéléromètres, de spires de flux...), les capteurs à utiliser sont ceux classiquement nécessaires à la commande électrique de la machine, à savoir les capteurs de courants et tensions.
- Au niveau de la méthodologie de diagnostic, ce travail s'inscrit dans le cadre des méthodes sans modèle, autrement qualifiées de "méthodes sans connaissance a priori". L'objectif de la RDF est de classer des formes mesurées en les comparant à des formes types. Une forme est une observation réalisée sur le processus, c'est à dire l'ensemble des  $d$  mesures effectuées à un instant donné. Une telle forme, qu'on appelle aussi échantillon, est représentée par un point dans un espace multidimensionnel de  $d$  paramètres. Les formes types sont en réalité des ensembles de points occupant des zones géométriques de l'espace de représentation appelées classes, où sont regroupées des formes semblables. En terme de diagnostic, les classes correspondent aux modes de fonctionnement connus et le fait de classer une nouvelle observation revient à identifier un de ces modes. Cela sous-entend que, non seulement un système de RDF possède dans sa base de données un ensemble d'apprentissage, mais qu'en plus, il est capable de prendre une décision pour toute nouvelle observation.

Dans ce contexte, le mémoire comporte trois chapitres. Le premier est consacré aux démarches classiques d'analyse de fonctionnement des moteurs asynchrones. Il comprend trois parties. La première partie fait état des différentes méthodes issues de modèles mathématiques et des résultats obtenus. La seconde partie présente comment, à partir des seules grandeurs mesurables, les informations relatives aux défauts de la machine peuvent être extraites. Enfin, la troisième partie est une synthèse des champs d'applications de ces méthodes.

---

Le deuxième chapitre traite des trois phases de l'élaboration d'un système de reconnaissance des formes en vue du diagnostic. La première phase, appelée étape d'apprentissage, consiste à définir le vecteur forme (la signature) ainsi que les classes, c'est à dire les modes de fonctionnement. En d'autres termes, il s'agit de construire la base de connaissance du système décisionnel. La phase d'apprentissage nécessite toute l'expérience humaine et la connaissance scientifique accumulées sur le processus pour mener à bien le choix des paramètres pertinents pour le diagnostic. Elle doit aboutir à la seconde phase consistant à élaborer une règle de décision qui devra être appliquée à toute nouvelle observation effectuée sur la machine, afin de la rattacher à l'une des classes existantes. Deux règles de décision sont ici proposées :

- La règle des  $k$  – plus proches voisins qui est une approche non paramétrique dans la mesure où elle réalise une estimation des lois de probabilité régissant les classes.
- Le calcul direct des frontières entre les classes. Cette deuxième règle de classification vise à calculer les coefficients des équations mathématiques des hyperplans qui délimitent les zones où sont situées les classes. Cette approche est donc analytique.

La troisième phase, couramment appelée phase d'exploitation, consiste à décider ou non d'affecter une nouvelle observation à l'une des classes connues, en vertu de la règle de décision.

Le troisième chapitre présente l'application expérimentale du système de diagnostic. Les défauts traités sont présentés dans la première partie de ce chapitre. Ce sont les cassures de barres et le déséquilibre d'alimentation, pour différents niveaux de charge en alimentation réseau et par ondulation de tension. La deuxième partie est consacrée aux méthodes d'analyses des signaux qui ont permis de définir le vecteur forme. Elle s'appuie sur la connaissance des défauts de la machine et des signatures permettant de les détecter et de les identifier. Les algorithmes de sélection de paramètres du vecteur forme utilisé par le système de décision et présentés au chapitre 2 (phase d'apprentissage), sont mis en application. Parmi l'ensemble des paramètres calculés, il s'agit de ne retenir que les plus efficaces en terme de pouvoir discriminant. La dernière partie consiste à définir la procédure de diagnostic par RDF appliquée à la machine asynchrone. Les deux règles de décision ( $k$ -PPV et frontières directes) sont alors comparées sur des essais sans et avec ondulation en présence ou non de défauts.



---

# CHAPITRE I – Synthèse des méthodes de diagnostic de la machine asynchrone

---

## *1.1 Introduction*

De nombreuses défaillances peuvent apparaître sur les machines asynchrones. Elles peuvent être électriques, mécaniques ou magnétiques. Leurs causes, très variées, sont classées en trois groupes [O'Donnell, 85] :

- les initiateurs de défauts : surchauffe du moteur, usures des éléments mécaniques (roulements à billes), rupture de fixations, problème d'isolation électrique, surtension transitoire...
- les contributeurs aux défauts : surcharge fréquente, température ambiante élevée, ventilation défaillante, humidité, fortes vibrations, vieillissement...
- les défauts sous jacents et erreurs humaines : défauts de fabrication, composants défectueux, protections inadaptées, absence de maintenance...

Ces différentes causes peuvent provoquer une dégradation des performances ou des arrêts intempestifs du système. Le diagnostic précoce des pannes est donc nécessaire.

De nombreuses méthodes ont été mises au point. La plupart sont basées sur une analyse vibratoire. Nous nous intéressons essentiellement à celles basées sur les signaux électriques.

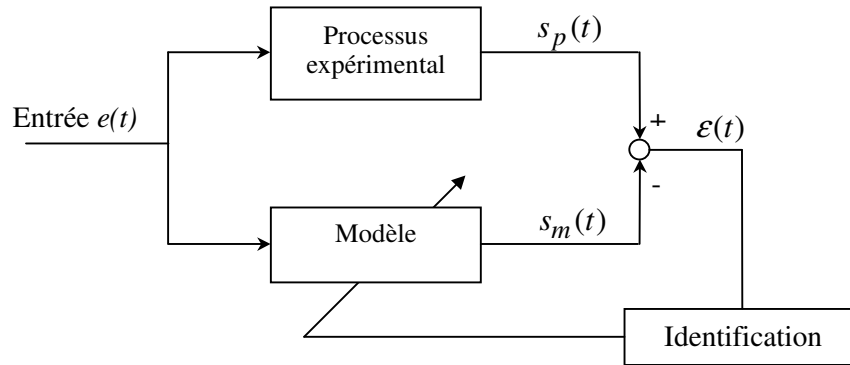
Le but de ce chapitre est de présenter un état de l'art des méthodes de diagnostic appliquées à la machine asynchrone. Ces méthodes sont classées en deux catégories, celles qui utilisent un modèle mathématique du système et celles qui consistent à analyser les signaux mesurés. Nous ferons une synthèse des champs d'application de ces méthodes.



## 1.2 Méthodes de diagnostic avec connaissance a priori

### 1.2.1 Principe

Ces méthodes de diagnostic sont généralement employées à partir d'une modélisation physique de la machine. Elles comparent l'évolution du modèle avec celle du processus physique. Considérons la machine asynchrone comme un processus expérimental dont il est question de rechercher, à chaque instant ou période d'échantillonnage, le modèle comportemental correspondant. La figure 1.1 traduit cette procédure.



**Fig. 1.1.** Principe du diagnostic par modèle.

On part d'un modèle du processus à surveiller sain. Si la sortie du modèle  $s_m(t)$  correspond à la sortie du processus  $s_p(t)$ , alors le modèle fournit une estimation des grandeurs caractéristiques du fonctionnement sans défaut. La détection de défaillance est réalisée par le suivi de l'erreur de sortie  $\varepsilon(t)$  (méthode des résidus) ou par la mise en évidence d'un écart entre le modèle (qui s'éloigne alors de la physique du phénomène) et le processus réel.

Dans le second cas, l'erreur de sortie peut être minimisée en modifiant les paramètres structuraux du modèle. De cette manière les paramètres inhérents au système sont identifiés et suivis même en cas de défaillance. Le modèle adapté donne alors une signature du mode de fonctionnement.

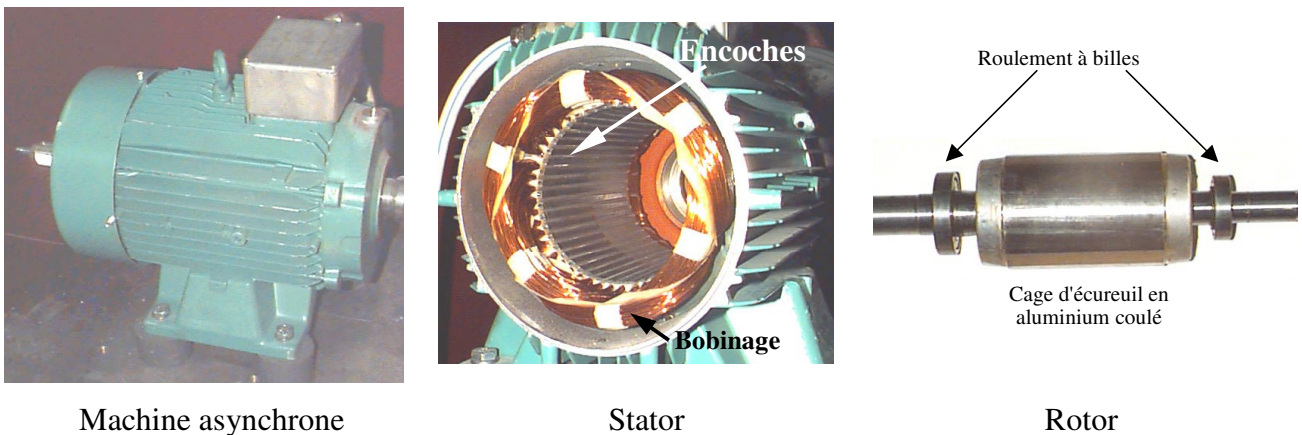
Nous présentons dans ce qui suit les modèles de comportement dynamique de la machine asynchrone destinés soit à la recherche des signatures des défauts, soit à l'analyse de fonctionnement en présence de défaut.

## 1.2.2 Modèle de la machine asynchrone

### 1.2.2.1 Transformation triphasée / diphasée

Les photographies suivantes (figure 1.2) présentent la machine asynchrone et ses deux principales composantes que sont le stator et le rotor.

La machine asynchrone est constituée de trois enroulements logés symétriquement dans les encoches du stator et de la cage conductrice intégrée au rotor. Ce dernier est assimilable à trois enroulements identiques parfaitement symétriques et en court-circuit.



**Fig. 1. 2.** Photo d'une machine asynchrone 5,5 kW avec le stator et le rotor qui la constituent

On schématise la disposition de ces enroulements (notés a, b, c) sur la figure 1.3 dans un plan perpendiculaire à l'axe de rotation.

On modélise la machine asynchrone à cage, en fonctionnement sain, dans le repère a, b, c selon les équations (1. 1) en émettant les hypothèses simplificatrices suivantes:

- Le circuit magnétique est linéaire, non saturé, ce qui signifie que les flux sont des fonctions linéaires des courants.
- L'induction dans l'entrefer a une répartition spatiale sinusoïdale.
- La densité surfacique des courants dans les conducteurs est uniforme.

Les équations électriques au stator et au rotor sont alors données par :

$$\begin{aligned} [V_s] &= R_s \cdot [I_s] + \frac{d[\Phi_s]}{dt} \\ [V_r] &= 0 = R_r \cdot [I_r] + \frac{d[\Phi_r]}{dt} \end{aligned} \quad (1. 1)$$

$$\text{Avec : } [\mathbf{I}_{s,r}] = \begin{bmatrix} i_{s,ra} \\ i_{s,rb} \\ i_{s,rc} \end{bmatrix}, [\mathbf{V}_{s,r}] = \begin{bmatrix} V_{s,ra} \\ V_{s,rb} \\ V_{s,rc} \end{bmatrix} \text{ et } [\Phi_{s,r}] = \begin{bmatrix} \varphi_{s,ra} \\ \varphi_{s,rb} \\ \varphi_{s,rc} \end{bmatrix}$$

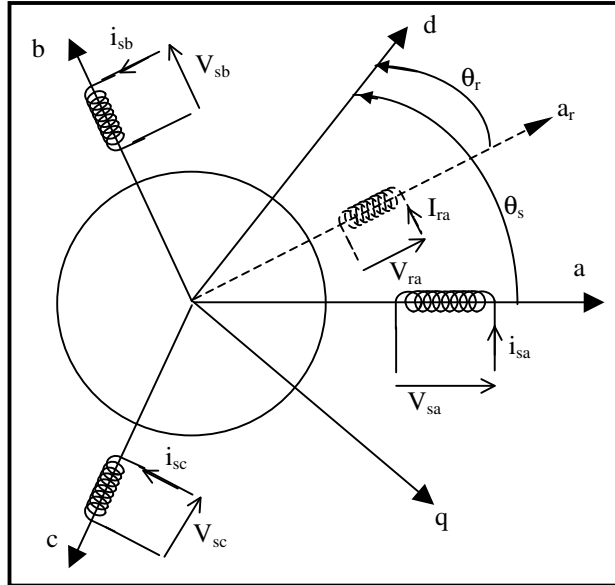


Fig. 1. 3. Représentation schématique des enroulements de la machine dans le plan a, b, c.

### 1.2.2.2 Transformation de Park

C'est la transformation des enroulements de phase a, b, c de la machine originelle en des enroulements disposés selon deux axes appelés d et q et équivalents du point de vue électrique et magnétique.

Une matrice unique de transformation est définie pour les courants, les tensions et les flux :

$$[\mathbf{T}] = \sqrt{\frac{2}{3}} \cdot \begin{bmatrix} \cos\theta_s & \cos(\theta_s - \frac{2\pi}{3}) & \cos(\theta_s - \frac{4\pi}{3}) \\ -\sin\theta_s & -\sin(\theta_s - \frac{2\pi}{3}) & -\sin(\theta_s - \frac{4\pi}{3}) \\ \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} \end{bmatrix} \quad (1.2)$$

Ainsi :

$$\begin{aligned}
 [i_{s,rdq0}] &= [T] \cdot [i_{s,abc}] \\
 [V_{s,rdq0}] &= [T] \cdot [V_{s,abc}] \\
 [\varphi_{s,rdq0}] &= [T] \cdot [\varphi_{s,abc}]
 \end{aligned} \tag{1.3}$$

$\theta_s$  est l'angle électrique entre le stator et l'axe d. L'indice 0 désigne les composantes homopolaires (ces composantes sont égales à zéro dans le cas d'un système triphasé équilibré).

Si  $\theta_s$  est nul alors on obtient la matrice [C], appelée matrice de Concordia telle que :

$$[C] = \sqrt{\frac{2}{3}} \cdot \begin{bmatrix} 1 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} & -\frac{\sqrt{3}}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{1.4}$$

Dans ce cas les axes d et q sont appelés  $\alpha$  et  $\beta$ .

Le coefficient  $\sqrt{\frac{2}{3}}$  est lié au changement de base et la puissance est conservée lors de cette transformation.

### 1.2.2.3. Choix du référentiel

Les équations de la machine peuvent être réécrites dans le repère d - q à condition de choisir le référentiel. Pour un référentiel immobile par rapport au stator, la convention utilisée est :

$$\frac{d\theta_s}{dt} = 0 \text{ et } \frac{d\theta_r}{dt} = \omega$$

il vient :

$$\text{Au stator : } \begin{cases} V_{sd} = R_s \cdot i_{sd} + \frac{d\varphi_{sd}}{dt} \\ V_{sq} = R_s \cdot i_{sq} + \frac{d\varphi_{sq}}{dt} \end{cases} \tag{1.5}$$

$$\text{Au rotor : } \begin{cases} V_{rd} = 0 = R_r \cdot i_{rd} + \frac{d\varphi_{rd}}{dt} - \omega \cdot \varphi_{rq} \\ V_{rq} = 0 = R_r \cdot i_{rq} + \frac{d\varphi_{rq}}{dt} + \omega \cdot \varphi_{rd} \end{cases} \tag{1.6}$$

D'autres référentiels peuvent être choisis :

- Référentiel lié au rotor avec :

$$\frac{d\theta_s}{dt} = \omega \text{ et } \frac{d\theta_r}{dt} = 0$$

- Référentiel lié au champ tournant avec :

$$\frac{d\theta_s}{dt} = \omega_s \text{ et } \frac{d\theta_r}{dt} = \omega_r$$

Les flux s'expriment en fonction des courants de la manière suivante :

$$\begin{bmatrix} \varphi_{sd} \\ \varphi_{rd} \end{bmatrix} = \begin{bmatrix} L_s & L_m \\ L_m & L_r \end{bmatrix} \cdot \begin{bmatrix} i_{sd} \\ i_{rd} \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} \varphi_{sq} \\ \varphi_{rq} \end{bmatrix} = \begin{bmatrix} L_s & L_m \\ L_m & L_r \end{bmatrix} \cdot \begin{bmatrix} i_{sq} \\ i_{rq} \end{bmatrix} \quad (1.7)$$

Le coefficient de fuite total et les constantes de temps statorique et rotorique sont définis respectivement par :

$$\sigma = 1 - \frac{L_m^2}{L_s \cdot L_r} \quad T_s = \frac{L_s}{R_s} \quad \text{et} \quad T_r = \frac{L_r}{R_r} \quad (1.8)$$

Pour tous les modèles de la machine décrits ci-dessus, l'expression du couple électromagnétique est donnée par l'équation (1.9) :

$$C_{em} = p \cdot (\varphi_{sd} \cdot i_{sq} - \varphi_{sq} \cdot i_{sd}) \quad (1.9)$$

Où  $p$  est le nombre de paires de pôles.

L'évolution de la vitesse de rotation en fonction du couple électromagnétique et de la charge de la machine caractérisée par le couple résistant  $C_r$ , est décrite par l'expression suivante (relation fondamentale de la dynamique) :

$$C_{em} - C_r = J \cdot \frac{d\omega}{dt} + f_t \cdot \omega \quad (1.10)$$

$J$  et  $f_t$  sont respectivement le moment d'inertie et le coefficient de frottement visqueux de la machine.

### 1.2.3. Application au diagnostic

Dans la partie précédente, nous avons présenté les équations régissant le fonctionnement de la machine asynchrone saine. Les paramètres de ces modèles sensibles aux défauts peuvent être identifiés ou observés [Schaeffer, 99 a] par des algorithmes spécifiques. En présence de défauts, ces paramètres dérivent et le modèle s'écarte alors de la physique du phénomène. Mais l'évolution des paramètres permet de caractériser le défaut.

### 1.2.3.1. Suivi du couple électromagnétique

En utilisant les systèmes d'équation en tension (1. 5) et (1. 6) appliqués au stator et au rotor, il est possible d'étudier le déséquilibre qu'entraîne la rupture de barres au rotor [Fiser, 96] par reconstitution de la distribution du champ magnétique et par l'étude de l'évolution du couple et des courants statoriques. A partir du modèle mathématique de la machine, on peut étudier l'influence de ce défaut sur la distribution du champ et sur les caractéristiques du couple et des courants statoriques.

Issus d'une simulation du fonctionnement à partir du modèle, les caractéristiques  $C(\omega)$  et  $I(\omega)$  (couple et courants en fonction de la vitesse) mettent en évidence les performances électromécaniques de la machine.

En écrivant l'équation d'état électrique et mécanique [Yahoui, 96] procède à l'estimation les composantes  $\varphi_{rd}$  et  $\varphi_{rq}$  du flux rotorique (notées  $\hat{\varphi}_{rd}$  et  $\hat{\varphi}_{rq}$ ) dans le but d'observer le couple électromagnétique. Les défauts étudiés dans leur article sont les ruptures de barres et d'anneaux. Dans un repère d'axe lié au stator l'équation d'état à la forme suivante :

$$\mathbf{U} = \mathbf{B} \cdot \mathbf{x} + \mathbf{A} \cdot \frac{d(\mathbf{x})}{dt} \quad (1. 11)$$

Le vecteur d'état  $\mathbf{x}$  est composé des courants statoriques et rotoriques dans les axes d et q et de la vitesse de rotation du rotor  $\omega$ .

Le vecteur  $\mathbf{U}$  ainsi que les matrices  $\mathbf{A}$  et  $\mathbf{B}$  sont respectivement étendus au couple résistant, au moment d'inertie et aux deux composantes du courant rotorique. L'équation (1.11) mise sous forme canonique donne:

$$\frac{d(\mathbf{x})}{dt} = \mathbf{A}_1 \cdot \mathbf{x} + \mathbf{B}_1 \cdot \mathbf{U} \quad (1. 12)$$

Avec  $\mathbf{A}_1 = (\mathbf{A})^{-1}(-\mathbf{B})$  et  $\mathbf{B}_1 = (\mathbf{A})^{-1}$

Nous pouvons obtenir une expression représentative du couple en fonction des composantes du courant statorique et du flux rotorique selon l'expression suivante :

$$\hat{C}_{em} = i_{sq} \cdot \hat{\varphi}_{rd} - i_{sd} \cdot \hat{\varphi}_{rq} \quad (1. 13)$$

Cette expression du couple observé est une image du couple réel qui, dans un repère d - q lié au stator, s'écrit :

$$C_{em} = p \frac{L_m}{L_r} (i_{sq} \cdot \varphi_{rd} - i_{sd} \cdot \varphi_{rq}) \quad (1. 14)$$

La dissymétrie électrique du rotor fait apparaître des harmoniques de dentures rotoriques, ce qui

modifie le champ d'entrefer. Le problème peut alors être détecté en analysant les fréquences d'encoche :

$$f_{sb} = f_s \cdot \left[ \frac{Nr \cdot (1-g)}{p} \pm n \right] \pm 2 \cdot g \cdot f_s \quad (1.15)$$

$f_{sb}$  : fréquence d'encoche au rotor

Nr : nombre d'encoche au rotor

$f_s$  : fréquence d'alimentation (fréquence des courants statoriques)

$n = 1, 2, 3, 4, \dots$

$g$  : le glissement

Ces fréquences sont donc identifiables en analysant le spectre du couple électromagnétique observé. Pour une alimentation à 50 Hz, [Yahoui, 95] porte l'attention sur la variation d'amplitude des raies à 600 et 900Hz en présence de barres cassées et des deux premiers harmoniques du couple qui augmentent en présence d'excentricité.

En s'intéressant aux excentricités statiques et dynamiques (voir plus loin dans le chapitre) du rotor, [Kral, 99] utilise le modèle triphasé pour estimer le flux statorique dans le plan complexe, ce qui permet d'estimer par la même occasion le couple utile. Ce dernier s'exprime en fonction des expressions complexes  $\underline{I}_s$  et  $\underline{\Phi}_s$  du courant et du flux statorique.

$$C_u = -\text{Im}(\underline{I}_s^* \cdot \underline{\Phi}_s) \quad (1.16)$$

$\underline{I}_s^*$  étant la forme conjuguée de l'expression complexe du courant.

En présence d'excentricité, le spectre du courant présente des raies aux fréquences :

$$f_{exc} = \left( 1 \pm n \frac{1-g}{p} \right) \cdot f_s \quad (1.17)$$

Le spectre du couple utile présente aussi des informations à ce sujet aux fréquences :

$$f_{Cu} = n \cdot \left( 1 - \frac{1+g}{p} \right) \cdot f_s \quad (1.18)$$

### 1.2.3.2. Identification des paramètres de la machine asynchrone

Le modèle de la machine est caractérisé par des paramètres électriques et mécaniques. A partir de l'acquisition des grandeurs électriques sur une machine asynchrone, l'identification de ces paramètres permet de construire un modèle fidèle à la machine réelle. En diagnostic, le suivi de ces paramètres permet de caractériser les défauts avec précision.

Deux types de paramètres sont à identifier :

- les paramètres mécaniques
- les paramètres électriques

#### 1.2.3.2.1 Équations d'état

A partir du modèle diphasé dans le repère  $\alpha, \beta$  lié au stator [Chouiter, 97], les cinq paramètres à identifier sont représentés par le vecteur  $[R_s \ R_r \ L_s \ L_r \ L_m]$ . Ces paramètres sont :

- $R_s$  : La résistance par phase du stator
- $R_r$  : La résistance par phase du rotor
- $L_s$  : L'inductance propre cyclique du stator
- $L_r$  : L'inductance propre cyclique du rotor
- $L_m$  : L'inductance mutuelle entre le stator et le rotor

Le modèle d'état de la machine peut être établi en choisissant le référentiel approprié (paragraphe 1.2.2.3).

L'équation d'état est la suivante :

$$\begin{cases} \frac{d\mathbf{x}(t)}{dt} = \mathbf{A}(\omega_r(t)) \cdot \mathbf{x}(t) + \mathbf{B} \cdot \mathbf{u}(t) \\ \mathbf{y}(t) = \mathbf{C} \cdot \mathbf{x}(t) \end{cases} \quad (1.19)$$

Dans la formulation (1.19),  $\mathbf{x}(t)$  représente le vecteur d'état,  $\mathbf{y}(t)$  le vecteur de sortie et  $\mathbf{u}(t)$  la commande. La matrice  $\mathbf{A}$  est appelée matrice dynamique ou matrice d'état,  $\mathbf{B}$  la matrice d'entrée ou de commande et  $\mathbf{C}$  la matrice de sortie.

$$\text{Avec } \mathbf{x}(t) = \begin{bmatrix} i_{s\alpha} \\ i_{s\beta} \\ \varphi_{r\alpha} \\ \varphi_{r\beta} \end{bmatrix} \text{ et } \frac{d(\mathbf{x}(t))}{dt} = \begin{bmatrix} i'_{s\alpha} \\ i'_{s\beta} \\ \varphi'_{r\alpha} \\ \varphi'_{r\beta} \end{bmatrix} \text{ on obtient :}$$



$$\begin{bmatrix} i'_{s\alpha} \\ i'_{s\beta} \\ \varphi'_{ra} \\ \varphi'_{r\beta} \end{bmatrix} = \begin{bmatrix} -\left[\frac{1}{T_s\sigma} + \frac{1-\sigma}{\sigma T_r}\right] & 0 & \frac{1-\sigma}{\sigma L_m T_r} & \frac{\omega_r(1-\sigma)}{\sigma L_m} \\ 0 & -\left[\frac{1}{T_s\sigma} + \frac{1-\sigma}{\sigma T_r}\right] & -\frac{\omega_r(1-\sigma)}{\sigma L_m} & \frac{1-\sigma}{\sigma L_m T_r} \\ \frac{L_m}{T_r} & 0 & -\frac{1}{T_r} & -\omega_r \\ 0 & \frac{L_m}{T_r} & \omega_r & \frac{1}{T_r} \end{bmatrix} \cdot \begin{bmatrix} i_{s\alpha} \\ i_{s\beta} \\ \varphi_{ra} \\ \varphi_{r\beta} \end{bmatrix} + \begin{bmatrix} \frac{1}{\sigma L_s} & 0 \\ 0 & \frac{1}{\sigma L_s} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} V_{s\alpha} \\ V_{s\beta} \end{bmatrix} \quad (1.20)$$

$$\mathbf{y}(t) = \mathbf{C} \cdot \mathbf{x}(t) \Rightarrow \begin{bmatrix} i_{s\alpha} \\ i_{s\beta} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} i_{s\alpha} \\ i_{s\beta} \\ \varphi_{ra} \\ \varphi_{r\beta} \end{bmatrix} \quad (1.21)$$

Il existe beaucoup d'autres représentations d'état de la machine. Les différences viennent du choix du référentiel et des variables d'état.

Par discrétisation on obtient :

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}(k)\mathbf{x}(k) + \mathbf{B}(k)\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) \end{aligned} \quad (1.22)$$

De ce fait, pour résoudre l'équation différentielle, on peut exprimer les matrices  $\mathbf{A}$  et  $\mathbf{B}$ , grâce à la matrice de transition  $\Phi(t)$  telle que :

$\Phi(t) = e^{\mathbf{A}t}$ , ce qui donne :

$$\begin{cases} \mathbf{A}(k) = \Phi(Te) = e^{\mathbf{A}Te} \approx \mathbf{I} + \mathbf{A}Te + \frac{(\mathbf{A}Te)^2}{2} \\ \mathbf{B}(k) = \int_0^{Te} e^{\mathbf{A}\tau} \mathbf{B} \cdot d\tau \approx Te \cdot \left( \mathbf{I} + \frac{\mathbf{A}Te}{2} \right) \cdot \mathbf{B} \end{cases} \quad (1.23)$$

$T_e$  est la période d'échantillonnage. Les conditions initiales pouvant être obtenues par la mesure des courants.

### 1.2.3.2.2 Modèle pour l'identification

Dans cette section nous présentons succinctement, à titre d'exemple, une des méthodes d'identification des paramètres de la machine nécessaire à la construction de son modèle. C'est la méthode du modèle de référence.

La méthode du modèle de référence est une méthode déterministe, c'est à dire qu'elle ne tient pas compte des perturbations du système [Lara, 95]. Cette méthode consiste à comparer les signaux issus du modèle du processus aux signaux mesurés. Le schéma de principe est représenté sur la figure 1.4.

Les entrées  $\mathbf{u}$  et les sorties  $\mathbf{y}$  sont issues des mesures. On estime le vecteur de paramètres  $\hat{\boldsymbol{\theta}}$  par minimisation du critère quadratique  $J(\boldsymbol{\theta})$ . Ce critère, dont l'expression est donnée par la relation (1.24), est fonction de l'écart entre la sortie mesurée  $\mathbf{y}$  et la sortie estimée  $\hat{\mathbf{y}}_m$  issue du modèle.

$$J(\boldsymbol{\theta}) = \sum_{k=1}^N \boldsymbol{\varepsilon}^2(k) = \sum_{k=1}^N [\mathbf{y}(k) - \hat{\mathbf{y}}_m(k)]^2 \quad (1.24)$$

Le processus et le modèle sont tous deux commandés par l'entrée  $\mathbf{u}$ .

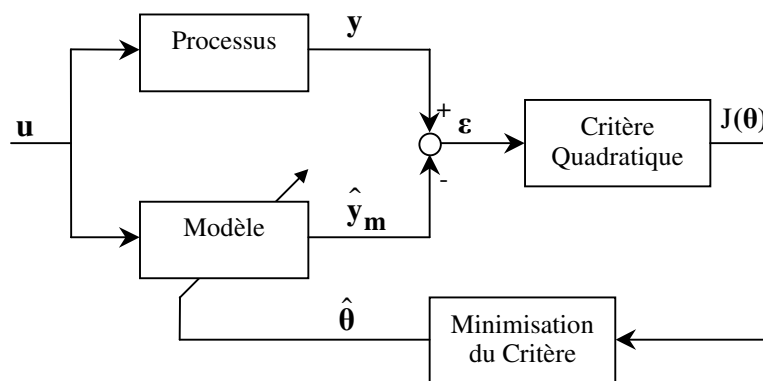


Fig. 1.4. Schéma de principe de la méthode du modèle de référence

Différents algorithmes de minimisation peuvent être utilisés (gradient, Newton – Raphson, Levenbert – Marquardt).

### 1.2.3.3. Paramètres représentatifs des défauts

Les paramètres caractéristiques de la machine asynchrone peuvent servir de très bons indicateurs de défaut. Nous présentons dans ce qui suit, à partir des informations recueillies dans la littérature, ceux qui sont représentatifs des défauts rencontrés au niveau de la machine.

L'une des défaillances les plus couramment étudiées dans la machine asynchrone à cage d'écuréuil est la rupture des barres rotoriques. Bien qu'elle soit identifiable à partir de l'analyse spectrale des courants statoriques on sait maintenant, grâce à l'estimation et à la surveillance de

l'évolution d'autres paramètres, détecter l'occurrence de ce défaut.

En effet, l'estimation par filtre de Kalman étendu [Saïd, 00] ou par observateur grand gain non linéaire [Boumégoura, 99] de la résistance  $R_r$  et du flux rotor  $\Phi_r$ , permet de voir que ceux-ci sont très sensibles aux défauts du rotor. La variation de  $R_r$  est d'autant plus importante que le nombre de barres cassées augmente.

Toutefois, le seul paramètre  $R_r$  ne permet pas de détecter la présence d'une ou plusieurs barres défectueuses. La valeur de cette résistance est liée à la variation de température du rotor et donc il suffirait que n'importe quel défaut pouvant entraîner une élévation de température apparaisse pour qu'elle varie. Cette ambiguïté persiste même si la valeur estimée est comparée à un seuil défini pour un fonctionnement à température ambiante.

L'expression d'une résistance en fonction de la température est donnée par :

$$R = R_0(1 + \alpha \cdot \Delta T) \quad (1.25)$$

Avec :

- $R_0$  : la valeur de  $R$  à  $T_0 = 25^\circ\text{C}$
- $\alpha$  : le coefficient de température de la résistance
- $\Delta T$  : la variation de la température

Ainsi, estimer un jeu de paramètres peut servir non seulement à mieux interpréter les observations mais aussi à élargir le champ d'investigation à d'autres types de défaut.

Dans son étude, [Moreau, 99] utilise l'identification des paramètres issus du modèle de Park pour détecter l'occurrence d'un défaut, suivi d'analyses fréquentielles pour distinguer les défauts du stator des défauts du rotor. Enfin l'estimation des paramètres à partir du modèle triphasé de la machine lui permet de reconnaître la phase statorique en défaut.

Ces études portent sur les défauts tels que les barres cassées et la réduction de spires ou les courts-circuits dans les enroulements statoriques. Les paramètres estimés sont les suivants :

- Pour le modèle de Park :  $\theta = [R_s \ R_r \ L_{fs} \ L_r]$
- Pour le modèle triphasé :  $\theta' = [R_{sa} \ R_{sb} \ R_{sc} \ L_{fsa} \ L_{fsb} \ L_{fsc} \ R_r \ L_{rp}]$

Les paramètres issus du modèle triphasé sont estimés à partir des paramètres du modèle de Park de la manière suivante :

$$R_{s3} \approx R_{sp}, \quad R_{r3} \approx R_{rp} \quad \text{et} \quad L_{rp} \approx \frac{3}{2} L_{rp3}, \quad \text{les indices 3 et p étant mis respectivement pour}$$

"triphasé" et "Park".

En terme de détection et de localisation des défauts, l'auteur tire les enseignements suivants :

- Un défaut au stator entraîne l'augmentation de  $R_s$  et la diminution de  $R_r$

- Un défaut au rotor entraîne la diminution de  $R_s$  et l'augmentation de  $R_r$

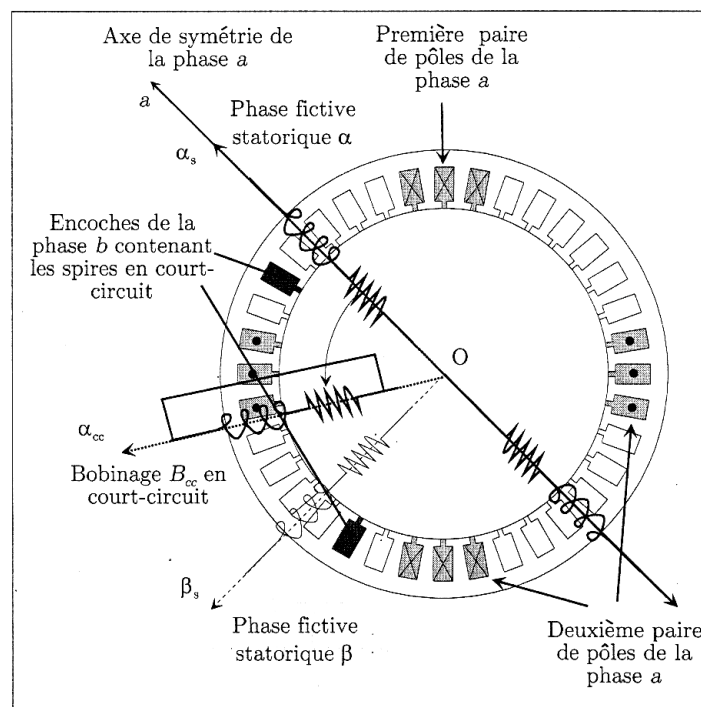
L'identification est confirmée par l'analyse spectrale des résidus de comparaison entre les paramètres réels et estimés:

- L'apparition d'une raie à  $f = n \cdot f_s$  est causée par un défaut au stator ( $n \geq 1$ )
- L'apparition d'une raie à  $f \ll f_s$  est causée par un défaut au rotor

Sachant que les fréquences fondamentales des courants statoriques et rotoriques sont respectivement égales à  $f_s$  et  $g \cdot f_s$  (où  $g$  est le glissement).

Dans le cas d'un défaut au stator (réduction de spires ou court-circuit) tous les paramètres issus du modèle triphasé changent mais seulement deux sont rigoureusement corrélés avec l'origine du défaut. Par exemple pour une réduction du nombre de spires sur la phase b : il y a une diminution de  $R_{sb}$  et  $L_{fsb}$  alors que les couples  $(R_{sa}, R_{sc})$  et  $(L_{fsa}, L_{fsc})$  augmentent.

Pour détecter un nombre plus faible de spires en court-circuit (jusqu'à 0,7% du nombre total de spire par phase), [Schaeffer, 99 b] fait une approche de type circuit qui met en évidence le fait qu'un court-circuit se traduit non seulement par la création d'une bobine fictive supplémentaire notée  $B_{cc}$  (figure 1.5) mais aussi par la diminution du nombre de spires dans la phase concernée.



**Fig. 1.5.** Représentation schématique des encoches stator [Schaeffer, 99 a]. Localisation des spires en court – circuit. La bobine fictive  $B_{cc}$  est représentée dans le repère  $\alpha, \beta$

Ainsi au modèle usuel de Park vient s'ajouter le paramètre  $\alpha$  (à ne pas confondre avec l'axe  $\alpha$  issu de la transformation diphasée) qui est le rapport entre le nombre  $n_{cc}$  de spires en court-circuit et le nombre total  $n_s$  de spires en série par phase du stator.

$$\alpha = \frac{n_{cc}}{n_s}$$

A la bobine  $B_{cc}$  s'associent les grandeurs vectorielles (courant et flux)  $\mathbf{i}_{cc}(t)$  et  $\Phi_{cc}(t)$ . L'angle  $p\gamma$  est l'angle entre l'axe de  $B_{cc}$  et l'axe de référence du bobinage statorique.

L'expression (1. 26) de  $\Phi_{cc}(t)$  est obtenue en considérant le schéma équivalent monophasé par phase de la machine asynchrone, et en introduisant la notion de flux magnétisant  $\phi_{ms}$  et le flux de fuite  $\phi_{fs}$ .

$$\Phi_{cc}(t) = \sqrt{\frac{3}{2}} \mathbf{Q}(p\gamma) \left( \alpha l_{ms} \cdot \mathbf{i}_s(t) + \alpha L_m \mathbf{P}(p\theta) \cdot \mathbf{i}_r(t) + \left( l_{ms} + \frac{3}{2} l_{fs} \right) \cdot \alpha^2 \sqrt{\frac{2}{3}} \cdot \mathbf{i}_{cc}(t) \right) \quad (1. 26)$$

$$\text{où } \mathbf{Q}(p\gamma) = \begin{bmatrix} \cos(p\gamma)\cos(p\gamma) & \cos(p\gamma)\sin(p\gamma) \\ \cos(p\gamma)\sin(p\gamma) & \sin(p\gamma)\sin(p\gamma) \end{bmatrix}, \quad l_{ms} \text{ et } l_{fs} \text{ sont respectivement les}$$

inductances magnétisante et de fuite que l'on aurait en l'absence de défaut,  $\theta$  est l'angle mécanique entre le stator et le rotor et  $\mathbf{P}$  la matrice de rotation.

Le système d'équation suivant résume le comportement électrique de la machine :

$$\begin{cases} \mathbf{u}_s(t) = R_s \mathbf{i}_s(t) + \frac{d}{dt} (\phi_{ms}(t) + \phi_{fs}(t)) \\ 0 = -\frac{3}{2} \frac{R_s}{\alpha} \cdot \mathbf{i}'_{cc}(t) + \mathbf{Q}(p\gamma) \frac{d}{dt} \phi_{ms}(t) \\ 0 = -R_r \mathbf{i}'_r(t) + \frac{d}{dt} \phi_{ms}(t) - p\dot{\theta} \cdot \mathbf{P}\left(\frac{\pi}{2}\right) \cdot \phi_{ms}(t) \end{cases} \quad (1. 27)$$

$\mathbf{i}'_{cc}(t)$  et  $\mathbf{i}'_r(t)$  sont les courants de  $B_{cc}$  et du rotor ramené au primaire (en considérant le schéma équivalent monophasé par phase de la machine).

Pour la représentation d'état les paramètres estimés a, b, c et d sont :

$$a = \frac{R_s}{\sigma L_s} + \frac{1-\sigma}{\sigma T_r}, \quad b = \frac{L_s(1-\sigma)}{T_r}, \quad c = \frac{1}{\sigma L_s}, \quad d = \frac{1}{T_r} \text{ et } \alpha.$$

Ce qui donne donc le vecteur des cinq paramètres, noté  $\lambda_0 = [a, b, c, d, \alpha]$ . L'influence du défaut sur le comportement entrée/sortie de la machine vient du fait que, dans la représentation d'état, le vecteur de sortie est fonction de l'entrée de commande par l'intermédiaire de  $\alpha$ .

Avec  $\mathbf{Y}$  le vecteur des courants diphasés mesurés et  $\mathbf{U}$  le vecteur des tensions diphasées d'alimentation (obtenus par la transformation de Concordia), la représentation d'état est la suivante :

$$\begin{cases} \dot{\mathbf{X}} = \mathbf{A}(\omega, \lambda_0) \cdot \mathbf{X} + \mathbf{B}(\lambda_0) \cdot \mathbf{U} \\ \mathbf{Y} = \mathbf{C}(\lambda_0) \cdot \mathbf{X} + \frac{3}{2} \frac{\alpha}{R_s} \cdot \mathbf{P}(p\theta) \mathbf{Q}(p\gamma) \mathbf{P}(-p\theta) \mathbf{U} \end{cases} \quad (1.28)$$

Ainsi, ce paramètre évolue de manière plus importante que  $R_s$ ,  $L_s$ ,  $T_r$  et  $\sigma$ . Surtout lorsqu'il s'agit d'un nombre faible de spires en court-circuit. Dans [Schaeffer, 99 a], le terme caractérisant la relation entrée/sortie est simplifié et donc, dans l'équation d'état le vecteur de sortie  $\mathbf{Y}$  s'exprime de la manière suivante :

$$\mathbf{Y} = \mathbf{C}(\lambda_0) \cdot \mathbf{X} + \alpha \omega_s \frac{\mathbf{U}}{|\mathbf{U}|^2} \quad (1.29)$$

Le rapport entre le nombre spires en court-circuit et le nombre total de spires dans une phase étudié est compris entre 1.5% et 16%. Cependant, cette approche reste sensible à la présence de défauts rotor.

## 1.2.4 Conclusion

Dans cette partie nous avons présenté les modèles permettant d'identifier les paramètres électriques et mécaniques de la machine asynchrone ou de développer des observateurs étendus permettant de suivre en ligne ces grandeurs.

Le tableau 1.1 présente les résultats obtenus à partir de l'état de l'art du diagnostic de la machine asynchrone par des méthodes basées sur la connaissance a priori de son fonctionnement. Ces résultats sont en réalité les signatures fréquentielles et "paramétriques" qui seraient caractéristiques des défauts étudiés.

Type de défaut	Signatures fréquentielles	Evolution des paramètres	Observations
Rupture de barres Rupture de portions d'anneaux	<p>Courants :</p> $f = f_s \cdot (1 \pm 2g)$ $f_b = f_s \cdot \left[ \frac{n}{p} \cdot (1 - g) \pm g \right]$ $f_{sb} = f_s \cdot \left[ \frac{N_r(1 - g)}{p} \pm n \right] \pm 2 \cdot g \cdot f_s$ <p>Résidus :</p> $f^* \ll f_s$	<p>DETECTION</p> <p><math>R_s</math> <i>diminue</i></p> <p><math>R_r</math> <u>augmente</u></p> <p><math>L_r</math> <i>diminue</i></p> <p><math>L_s</math> <i>diminue</i></p>	<p><i>Paramètres issus du modèle de Park</i></p> <p><i>*Fréquences des résidus de comparaison entre paramètres réels et estimés</i></p>
	Excentricité	<p>Courants :</p> $f_e = f_s \cdot \left[ 1 \pm n \frac{1 - g}{p} \right]$ <p>Couple utile :</p> $f_{Cu} = n \cdot \left[ 1 - \frac{1 + g}{p} \right] \cdot f_s$	
Court-circuit	<p>Résidus :</p> $f^* = n \cdot f_s$ <p>avec <math>n = 1, 2, 3 \dots</math></p>	<p>DETECTION</p> <p><math>R_s</math> <u>augmente</u></p> <p><math>R_r</math> <i>diminue</i></p> <p><math>L_r</math> <i>diminue</i></p> <p><math>L_f</math> <i>diminue</i></p> <p>LOCALISATION</p> <p><math>R_{sa}</math> et <math>L_{sa}</math> <u>augmentent</u> si défaut sur la phase a</p>	<p><i>*Fréquences des résidus de comparaison entre paramètres réels et estimés</i></p> <p><i>Rsa et Lsa issus du modèle triphasé</i></p>

**Tableau. 1. 1.** Premier récapitulatif de signatures

Le modèle découlant de la transformation des grandeurs électriques triphasées en composantes monophasées est la base, en analyse de fonctionnement ou en commande, du développement d'algorithmes d'observation ou d'estimation. Observer les grandeurs difficilement mesurables permet de réduire le nombre de capteur.

Les paramètres résistances, inductances, inductances mutuelles et la vitesse peuvent servir de bons indicateurs de défaut dès lors qu'ils sont correctement estimés.

---

Nous allons compléter cette étude en introduisant d'autres signaux (flux axial, vibrations mécaniques, signaux d'arbres, températures, et même rayonnements électromagnétiques). Ces signaux sont souvent analysés à partir de mesures directes (capteurs) ; ce qui introduit l'autre grande famille des méthodes de diagnostic, à savoir les méthodes de diagnostic sans modèle.



## 1.3 Méthodes de diagnostic sans connaissance a priori

### 1.3.1 Principe

Nous avons vu que la première famille de méthodes de diagnostic nécessite la connaissance du comportement dynamique de la machine asynchrone. Toutefois nous avons remarqué que les modèles ne permettaient pas d'étudier tous les défauts de la machine (défauts des roulements par exemple). Dans cette deuxième partie nous nous intéressons donc aux résultats obtenus à partir du suivi direct des grandeurs telles que les courants, les vibrations, les flux et le couple.

### 1.3.2 Suivi des grandeurs mesurables

#### 1.3.2.1 Analyse des courants

Les procédures de surveillance de fonctionnement sont généralement orientées vers des défaillances spécifiques intervenant sur l'une des trois parties de la machine : le stator, le rotor ou les paliers. Ces derniers supportent l'arbre moteur.

Pour extraire de manière précise les informations relatives aux défauts les recherches ont été particulièrement dirigées vers le spectre des courants statoriques pour deux raisons. Les courants sont faciles à mesurer, ils fournissent des informations sur de nombreux défauts.

Mais la mesure des signaux puis leurs traitements dans le domaine spectral ne peuvent servir, en terme d'identification, que si les composantes fréquentielles définies pour chaque défaut [Benbouzid, 99] [Schoen, 95 a] sont connues. Cette méthode ne s'applique actuellement qu'en régime permanent.

##### 1.3.2.1.1 Excentricité

La variation de la longueur de l'entrefer entraîne des variations dans la densité du flux d'entrefer. Les enroulements statoriques sont électriquement affectés. On retrouve alors ces effets dans le spectre des courants :

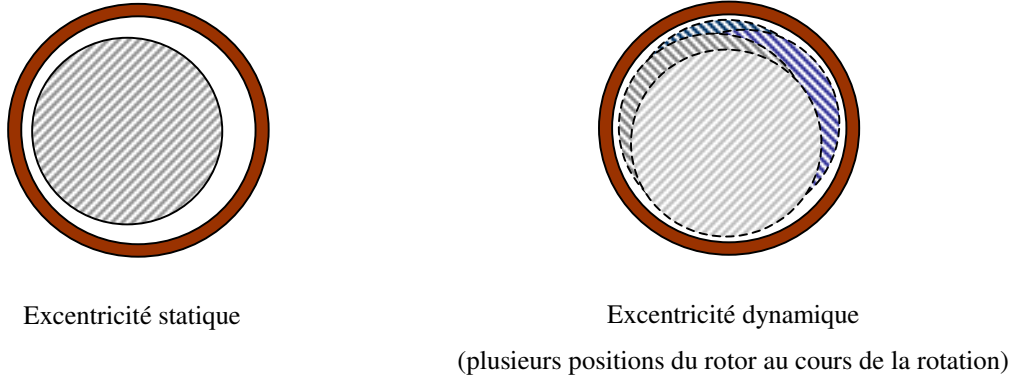
$$f_{exc} = f_s \cdot \left[ 1 \pm n \frac{1-g}{p} \right] \quad (1.30)$$

On distingue généralement trois cas d'excentricité [Sobczyk, 98] :

- l'excentricité statique, le rotor est déplacé du centre de l'alésage stator mais tourne toujours autour de son axe

- l'excentricité dynamique, le rotor est positionné au centre de l'alésage mais ne tourne plus autour de son axe
- l'excentricité qu'on pourrait qualifier de globale, associant les deux cas précédemment cités

On peut représenter l'excentricité statique et dynamique de la manière suivant :



**Fig. 1. 6.** Représentation de l'excentricité statique et dynamique

Les causes peuvent être le positionnement incorrect des paliers lors de l'assemblage, l'usure des paliers ou la torsion de l'arbre. Dans la référence [Benbouzid, 99], les auteurs utilisent des méthodes d'intelligence artificielle pour le diagnostic du défaut. Mais le plus important est l'identification du type d'excentricité qui est possible par la visualisation, dans le spectre des courants de phase, des composantes fréquentielles suivantes :

Pour l'excentricité statique :

$$f_{exc} = f_s \cdot \left[ 1 \pm \frac{n \cdot N_r (1 - g)}{p} \right] \tag{1. 31}$$

Pour l'excentricité dynamique :

$$f_{exc} = f_s \cdot [1 \pm 2 \cdot n(1 - g)] \tag{1. 32}$$

Pour le désalignement :

$$f_{exc} = f_s \cdot \left[ 1 \pm \frac{n \cdot (1 - g)}{p} \right] \tag{1. 33}$$

L'origine du désalignement est introduite dans le paragraphe suivant.

### 1.3.2.1.2 Défaillance des paliers

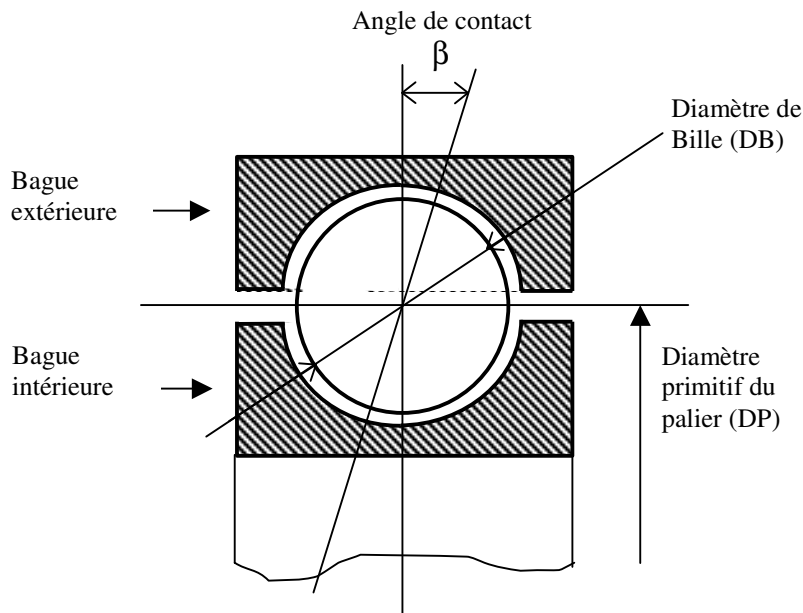
Lors de son installation, le palier est souvent appuyé sur l'arbre ou dans le carter. Il s'ensuit alors un possible désalignement qui peut apparaître de quatre manières différentes :

- le désalignement proprement dit où les deux paliers (supportant le même arbre) ne sont pas dans le même axe.
- la flexion de l'arbre

- l'inclinaison d'une bague extérieure de roulement
- l'inclinaison d'une bague intérieure de roulement

On peut retrouver les phénomènes vibratoires des paliers sur le spectre des courants statoriques en considérant que l'excentricité ou le désalignement du rotor conduit nécessairement à une variation de la densité de flux d'entrefer.

Etant donné que les paliers supportent le rotor, grâce aux roulements, un défaut de ceux-ci produira un mouvement radial entre le rotor et le stator. Les fréquences caractéristiques du défaut sont déduites des dimensions des paliers (DB et DP) et des roulements à billes, comme le montre la figure 1.7.



**Fig. 1. 7.** Vue en coupe de la partie supérieure d'un roulement à bille [Benbouzid, 99] [Schoen, 95 a] [Schoen, 95 b]

En fonction du nombre de billes noté  $n_b$ , on définit les composantes fréquentielles spécifiques aux vibrations des paliers notées  $f_{pal}$ . Ces composantes sont définies par la relation (1. 34)

$$f_{pal} = f_s \pm n \cdot f_{i,e} \tag{1. 34}$$

$n = 1, 2, 3 \dots$  "i , e" : interne, externe (bagues)

avec

$$f_{i,e} = \frac{n_b}{2} \cdot f_r \cdot \left[ 1 \pm \frac{DB}{DP} \cdot \cos \beta \right] \tag{1. 35}$$

Où  $f_r$  est la fréquence de rotation du rotor.

De manière plus précise, on définit les fréquences des bagues internes et externes de roulement

respectivement par les expressions (1. 36) et (1. 37)

$$f_e = \frac{n_b}{2} \cdot f_r \cdot \left[ 1 - \frac{DB}{DP} \cdot \cos \beta \right] \quad (1. 36)$$

$$f_i = \frac{n_b}{2} \cdot f_r \cdot \left[ 1 + \frac{DB}{DP} \cdot \cos \beta \right] \quad (1. 37)$$

Comme nous l'avons précisé, ces informations sont fonction des données du constructeur. On peut toutefois approcher les expressions de  $f_e$  et  $f_i$  pour un nombre faible de billes dans les roulements [Benbouzid, 99] par :

$$\begin{aligned} f_e &= 0.4 \cdot n_b \cdot f_r \\ f_i &= 0.6 \cdot n_b \cdot f_r \end{aligned} \quad (1. 38)$$

### 1.3.2.1.3 Ruptures de barres ou de portions d'anneaux

Nous n'avons pas jugé utile ici de mener notre étude sur la détection de barres cassées par la mesure des courants statoriques. En effet, il est bien connu qu'un traitement approprié de ces signaux fait apparaître des composantes dans leurs spectres aux fréquences  $f = f_s(1 \pm 2 \cdot g)$  [Thomson, 00].

En effet, lorsqu'une barre est cassée, des harmoniques de flux correspondant sont produits et induisent des harmoniques de courant dans l'enroulement statorique aux fréquences :

$$f = f_s(1 \pm 2 \cdot g) \quad (1. 39)$$

Par ailleurs ce défaut à aussi une influence sur les harmoniques de rang faible des courants statoriques :

$$f_b = f_s \cdot \left[ \frac{n}{p} \cdot (1 - g) \pm g \right] \quad (1. 40)$$

De plus les amplitudes des raies dans ces bandes latérales du fondamental augmentent avec le nombre de barres défailantes [Stein, 88]

## 1.3.2.2 Analyse des flux

La conversion électromécanique de l'énergie est localisée dans l'entrefer. Cette conversion est donc affectée par tout déséquilibre magnétique, mécanique, électrique ou électromagnétique au rotor ou au stator. Le flux d'entrefer et par-là même l'induction magnétique dans cet entrefer, le flux embrassé dans les enroulements statoriques ou encore le flux de fuite dans l'axe du rotor sont des

paramètres qui, à cause de leur sensibilité à tout déséquilibre de la machine, méritent d'être analysés.

Ces grandeurs sont difficilement mesurables mais des études ont été menées pour extraire des signatures spécifiques à certains défauts.

Le flux axial est toujours présent dans les machines électriques à cause des dissymétries inhérentes à leur fabrication. L'étude des variations de flux peut donc être une solution pour détecter et localiser un défaut au travers de l'utilisation de bobines exploratrices placées à l'extérieur de la machine, perpendiculairement à l'axe du rotor.

Dans l'étude [Hénao, 99] l'auteur met en évidence les composantes fréquentielles du flux de fuite dans l'axe du rotor qui permettent de détecter et d'identifier des courts-circuits dans les enroulements statoriques. L'étude théorique est reprise dans ce qui suit.

Le flux axial est le résultat de l'effet des courants sur les extrémités de la machine (têtes de bobines, anneaux de court-circuit). Dans une machine symétrique alimentée par un système de tensions équilibrées, par décomposition en séries de Fourier, l'induction magnétique d'entrefer  $B_{es}$  peut s'exprimer de la manière suivante :

$$B_{es}(\theta, t) = \sum_i B_i \cos(\omega_s t - ip\theta) \quad (1.41)$$

$B_i$  est l'amplitude de l'harmonique de rang  $(6n \pm 1)$ ,  $p$  le nombre de paires de pôles,  $\omega_s$  la pulsation des tensions statoriques et  $\theta$  l'angle mécanique associé à la position du rotor.

L'équation (1.41) est exprimée dans un repère lié au stator. Dans un repère lié au rotor, il suffit d'appliquer à  $\theta$  la relation suivante :

$$\theta = \theta_0 + \frac{\omega_s(1-g)}{p} \cdot t \quad (1.42)$$

La nouvelle expression spatio-temporelle de l'induction est alors :

$$\begin{aligned} B_{er}(\theta, t) = & B_1 \cos(g\omega_s t - p\theta_0) \\ & + B_5 \cos((6-5g)\omega_s t + 5p\theta_0) \\ & + B_7 \cos((7g-6)\omega_s t - 7p\theta_0) + \dots \end{aligned} \quad (1.43)$$

Cette expression donne implicitement les composantes harmoniques des courants au rotor.

### 1.3.2.2.1 Court - circuit entre spires au stator

Si on considère  $N_{cc}$  le nombre de spires en court-circuit et  $i_{cc}(t)$  le courant instantané dans celles-ci, alors la force magnétomotrice pulsante  $F_{cc}$  associée à un enroulement présentant des spires en

court-circuit peut être exprimée par :

$$F_{cc}(\theta, t) = \frac{1}{\pi} N_{cc} I_{cc} \sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{n\pi}{2p}\right) \cos(\omega_s t - \varphi \pm n\theta) \quad (1.44)$$

Avec :

$$i_{cc}(t) = I_{cc} \cos(\omega_s t - \varphi) \quad (1.45)$$

Cela nous amène à l'expression, dans le repère lié au rotor, de l'induction d'entrefer permettant d'identifier les fréquences intéressantes pour la détection de court-circuit à partir du flux axial.

$$B_{er,cc} = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} B_{ncc} \cos\left[\left(k \pm n \frac{1-g}{p}\right) \omega_s t - \varphi \pm n\theta_0\right] \quad (1.46)$$

$B_{ncc}$  est l'amplitude du  $n^{i\grave{e}me}$  harmonique,  $k$  un entier associé aux harmoniques de l'alimentation. Et  $n$  est un entier différent du nombre de pôles de la machine sinon la force magnétomotrice (1.43) est égale à zéro.

Ainsi ces fréquences caractéristiques sont [Hénao, 99] :

$$k \cdot f_s \pm n \cdot f_r \quad (1.47)$$

$$n \in \mathbb{N}$$

$$k \in \mathbb{N} \text{ et } k \text{ impair}$$

Toutes ces composantes, déjà présentes dans le spectre du flux pour la machine saine, vont augmenter en présence du défaut [Cabanais, 98]. Par ailleurs ces augmentations sont indépendantes du niveau de charge.

Les mêmes constatations ont été faites pour les composantes du flux axial liées aux barres rotoriques et aux paliers. Les composantes à prendre en compte pour les barres sont répertoriées dans le tableau 1.2. Et pour ce qui est des paliers, les auteurs mettent l'accent sur les raies situées dans la plage de fréquence 600 Hz à 900 Hz pour une alimentation à 50 Hz.

$f_s$	$(5 - 4g)f_s$
$(2g - 1)f_s$	$(6g - 5)f_s$
$(3 - 2g)f_s$	$(7 - 6g)f_s$
$(4g - 3)f_s$	$(8g - 7)f_s$

**Tableau. 1. 2.** Fréquences associées au flux axial

**1.3.2.2 Cas d'une dissymétrie au rotor**

Le flux statorique, bien qu'il soit non mesurable (mais estimable par un observateur), est un paramètre à ne pas négliger pour la surveillance du fonctionnement de la machine. Pour l'identification du défaut de barres cassées au rotor, les composantes caractéristiques du flux statorique sont les mêmes que celles précédemment citées dans le tableau 1.2 [Deng, 93].

**1.3.2.3 Analyse du couple électromagnétique**

Le couple électromagnétique peut être reconstruit à partir de deux des trois courants statoriques et par ajout de spires de flux dans ces deux phases (cela permet d'obtenir les forces électromotrices qui y sont induites).

Dans un repère d,q lié au stator, le couple électromagnétique a pour expression :

$$C_{em} = p \cdot (\Phi_{sd} \cdot i_{sq} - \Phi_{sq} \cdot i_{sd}) \tag{1.48}$$

Si on considère les courants et les flux dans un repère triphasé :







$$C_{em} = \frac{p}{\sqrt{3}} \cdot (\Phi_{s2} \cdot i_{s1} - \Phi_{s3} \cdot i_{s1} - \Phi_{s1} \cdot i_{s2} + \Phi_{s3} \cdot i_{s2} - \Phi_{s1} \cdot i_{s3} - \Phi_{s2} \cdot i_{s3}) \tag{1.49}$$

$\Phi_{s1}$ ,  $\Phi_{s2}$  et  $\Phi_{s3}$  sont les flux embrassés sous un pôle pour les phases 1, 2 et 3 respectivement et  $i_{s1}$ ,  $i_{s2}$  et  $i_{s3}$  sont les courants de phase au stator.

Avec une alimentation triphasée sans neutre on a  $i_{s1} = -i_{s2} - i_{s3}$  et par conséquent,  $\Phi_{s1} = -\Phi_{s2} - \Phi_{s3}$ . Ceci permet d'obtenir :

$$C_{em} = \sqrt{3} p \cdot (\Phi_{s3} \cdot i_{s2} - \Phi_{s2} \cdot i_{s3}) \tag{1.50}$$

En se basant sur ces signaux [Thollon, 93] analyse le couple électromagnétique d'une machine asynchrone présentant des portions d'anneaux cassées. Nous présentons dans le tableau 3 les résultats obtenus concernant les variations des harmoniques du couple en présence du défaut.

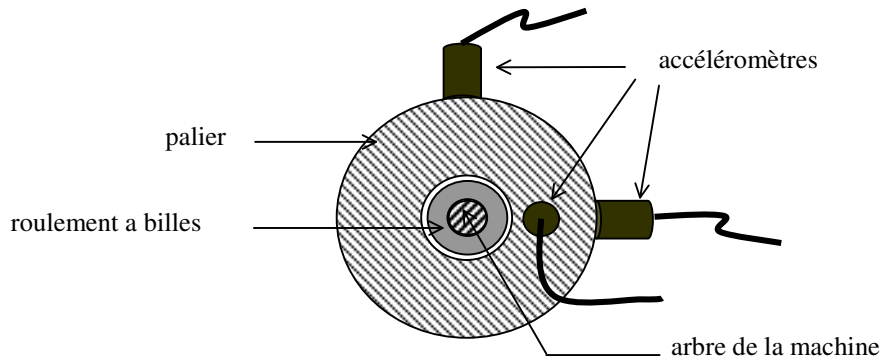
50 Hz	100 Hz	200 Hz	300 Hz	700 Hz	800 Hz
					

**Tableau. 1. 3.** Variations des harmoniques du couple électromagnétique cas de ruptures d'anneaux au rotor (fondamental à 50 Hz) [Thollon, 93]

L'harmonique à 700 Hz représente les composantes fréquentielles dues au nombre de barres et d'encoches au rotor. Leur expression est donnée par la relation 1.15 dans la première partie.

### 1.3.2.4. Analyses vibratoires de la machine

L'analyse vibratoire de la machine asynchrone peut être menée grâce à des accéléromètres placés sur les paliers dans les directions verticales, axiales et radiales.



**Fig. 1. 8.** Mesures vibratoires sur la machine asynchrone

Au même titre que les courants, les vibrations de la machine fournissent des informations sur pratiquement tous les défauts mécaniques fréquemment rencontrés. Plutôt que de disposer d'une étude au cas par cas, nous avons réalisé une synthèse de l'ensemble des informations que peuvent fournir ces signaux en fonction des défauts, à partir de la référence [Calis, 99], qui présente une comparaison entre les signatures des courants et les signatures vibratoires. Ces observations sont résumées dans le tableau 1.3.

Nous avons remarqué que les composantes fréquentielles, qui permettent d'identifier les défauts cités, dépendent de la fréquence de rotation du rotor. De ce fait, ces signatures ne sont pas assez distinctes les unes par rapport aux autres. Leur utilisation poserait des problèmes de discrimination des types de défaut. Par contre, pour la détection d'un défaut ou d'un déséquilibre au rotor, les vibrations de la machine s'avèrent être des signaux pertinents.



Type de défaut	Analyse vibratoire	Observations
Rupture de barres Rupture de portions d'anneaux	$f_r \pm 2 \cdot g \cdot f_s$	Les vibrations augmentent dans l'axe du rotor
Excentricité Statique Dynamique	Exc. Statique : $2 \cdot f_s$ ou des composantes à $k \cdot f_r \pm f_s$ Exc. Dynamique : $f_r \pm 2 \cdot g \cdot f_s$ ou $f_r$	L'excentricité dynamique est fonction du temps et de l'espace L'excentricité statique est fonction de l'espace seulement
Désalignement	$f_r$ et $2 \cdot f_r$ principalement	On parle de désalignement : Parallèle et/ou angulaire Des paliers De l'arbre (torsion)
Déséquilibre mécanique	$f_r$	Au niveau de la charge notamment Vibrations plus importantes dans la direction radiale
Défaillance des paliers	$n \cdot f_e$ $f_i \pm k \cdot f_r$	Bagues externes Bagues internes Ces composantes fréquentielles sont modulées par la vitesse

**Tableau. 1. 4.** *Analyses vibratoires, signatures fréquentielles [Calis, 99]*

## 1.4 Synthèse générale des résultats

### 1.4.1 Champs d'application des méthodes de diagnostic

Cette partie présente une synthèse des signatures des défauts. Nous présentons dans ce qui suit les défauts étudiés à partir des méthodes à base de connaissances a priori et des méthodes de diagnostic sans connaissance a priori.

- Méthodes de diagnostic à base de connaissance a priori

Les défauts étudiés sont :

- Ruptures de barres ou de portions d'anneaux au rotor
- Excentricité statique et dynamique
- Réduction de spires ou court-circuit entre spires au stator

- Méthodes de diagnostic sans connaissance a priori

Les défauts étudiés sont :

- Ruptures de barres ou de portions d'anneaux au rotor
- Excentricité statique et dynamique
- Défaillance des paliers :
  - ✓ désalignement de l'arbre
  - ✓ flexion de l'arbre
  - ✓ défauts sur les bagues externes de roulements
  - ✓ défauts sur les bagues internes de roulements
- Court-circuit entre spires au stator

Enfin nous reprenons dans les tableaux suivants les signaux ainsi que les paramètres de la machine asynchrone qui méritent d'être pris en compte pour détecter, identifier ou localiser les défauts précédemment cités.

Type de défaut	Signaux et signatures				Variation des paramètres	Méthode de diagnostic		Observations
	Courants	Couple	Vibrations	Flux		A M	S M	
Rupture de barres ou de portions d'anneaux	$f = f_s \cdot (1 \pm 2g)$ $F_{enc} \pm 2n f_s \pm 2g f_s$	Fréq. d'encoches sur Cem : $f_{sb} = f_s \cdot \left[ \frac{N_r(1-g)}{p} \pm n \right] \pm 2 \cdot g \cdot f_s$	$f_r \pm 2 \cdot g \cdot f_s$ $F_{enc} \pm 2n f_s \pm 2g f_s$	Flux axial : $f_s$ $(2g - 1)f_s$ $(3 - 2g)f_s$ $(4g - 3)f_s$ $(5 - 4g)f_s$ $(6g - 5)f_s$ $(7 - 6g)f_s$ $(8g - 7)f_s$	DETECTION Rs <i>diminue</i> Rr <i>augmente</i> Lr <i>diminue</i> Lf <i>diminue</i> $f^* \ll f_s$			AM : avec modèle SM : sans modèle  *Fréquences des résidus de comparaison entre paramètres réels et estimés  <b>Rs : résistance totale de l'enroulement stator</b>
Excentricité: - statique - dynamique - globale	Statique : $f_{exc} = f_s \cdot \left[ 1 \pm \frac{n \cdot N_r(1-g)}{p} \right]$ Dynamique : $f_{exc} = (N_r \cdot f_r + f_s) \pm f_r$ Globale : $f_{exc} = f_s \cdot \left[ 1 \pm n \frac{1-g}{p} \right]$	analyse du couple utile : $f_{Cu} = n \cdot \left[ 1 - \frac{1+g}{p} \right] \cdot f_s$	Exc. Statique : $2 \cdot f_s$ ou des composantes à $k \cdot f_r \pm f_s$ Exc. Dynamique : $f_r \pm 2 \cdot g \cdot f_s$ ou $f_r$					Augmentation des deux premiers harmoniques du couple électromagnétique observé
Méthode de diagnostic								
Avec Modèle								
Sans Modèle								

Tableau. 1. 5. Synthèse des signatures des défauts.

Type de défaut	Signaux et signatures				Variation des paramètres	Méthode de diagnostic		Observations
	Courants	Couple	Vibrations	Flux		A M	S M	
Défaillance des paliers	$f_{pal} = f_s \pm n \cdot f_{i,e}$ avec : $f_e = 0.4 \cdot n_b \cdot f_r$ $f_i = 0.6 \cdot n_b \cdot f_r$		Bagues externes : $n \cdot f_e$ bagues internes : $f_i \pm k \cdot f_r$				AM : avec modèle SM : sans modèle	
Désalignement	$f_s \pm k \cdot f_r$		$f_r$ et $2 \cdot f_r$ principalement				On parle de désalignement : <ul style="list-style-type: none"> <li>▪ Parallèle et/ou angulaire</li> <li>▪ Des paliers</li> <li>▪ De l'arbre (torsion)</li> </ul>	
Court-circuit entre spires au stator				Flux axial : $k \cdot f_s \pm n \cdot f_r$ Altération possible du flux stator	DETECTION Rs <i>augmente</i> Rr <i>diminue</i> Lr <i>diminue</i> Lf <i>diminue</i> $f^* = n \cdot f_s$ augmentation des paramètres du vecteur : $\lambda_0 = [a \ b \ c \ d \ \alpha]$		k = 1,3,5... n = 1,2,3... n ≠ 2p LOCALISATION Rsa et Lsa <i>augmentent</i> <i>si défaut sur la phase a</i> *Fréquences des résidus de comparaison entre paramètres réels et estimés a, b, c, d fonction des paramètres de Park	
Méthode de diagnostic								
Avec Modèle								
Sans Modèle								

Tableau. 1. 6. Synthèse des signatures des défauts (suite).

## 1.4.2 Conclusion

L'objectif de cette étude était d'établir un état de l'art des méthodes de diagnostic mises en œuvre et des résultats obtenus.

Dans un premier temps nous nous sommes intéressés aux méthodes de diagnostic à base de modèles. Si le comportement électrique et dynamique de la machine est parfaitement connu, les modèles utilisés peuvent fournir une estimation fiable des grandeurs difficilement mesurables. Les signaux et paramètres de sortie sont alors utilisés soit pour la commande, soit pour la surveillance et le diagnostic.

D'un autre côté, des mesures vibratoires, électriques et électromagnétiques rendent directement compte de l'état de la machine.

Toutefois ces deux solutions doivent être utilisées de manière complémentaire : la mesure des courants et des vibrations par exemple peut suffire pour une surveillance en ligne. Mais l'estimation des paramètres et des variables apporte de manière moins onéreuse des informations supplémentaires.

La liste des signatures des défauts est sans aucun doute non exhaustive, mais ces résultats serviront de point de départ pour la recherche de la signature nécessaire à la reconnaissance des formes. En effet, pour être appliquée au diagnostic la reconnaissance des formes (RDF) nécessite la connaissance de l'état de fonctionnement du système étudié par le biais du vecteur forme.

Un des principaux avantages de l'application de la RDF en diagnostic est de n'utiliser qu'une signature pour caractériser le fonctionnement du système étudié. A l'aide de la connaissance acquise sur les défauts et sur le fonctionnement de la machine, nous pourrions construire un vecteur de paramètres qui constituera la signature.



---

## CHAPITRE II – Reconnaissance des formes et diagnostic

---

### 2.1. Introduction

La Reconnaissance des Formes est un des nombreux aspects de l'intelligence artificielle ou plus exactement de l'intelligence calculatoire. A partir d'un ensemble de données ou d'informations apprises, elle offre la possibilité d'interpréter toute nouvelle observation (ou forme). Les observations déjà connues sont regroupées en classes, constituant des prototypes auxquels la nouvelle observation est comparée pour être identifiée. Les algorithmes utilisés permettent donc de classer des observations dont les propriétés ont varié par rapport à une observation type [Dubuisson, 90]. Il s'agit donc d'un outil qui a, entre autres, la capacité d'apprentissage.

Il existe deux types de reconnaissance des formes. Le premier, appelé reconnaissance des formes structurelle, utilise la grammaire. Les formes y sont essentiellement caractérisées par des propriétés grammaticales. On pourra en trouver des exemples d'application dans [Friedman, 99].

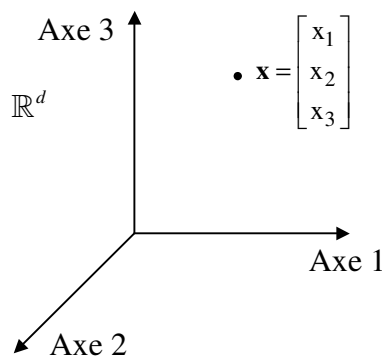
Le second type de reconnaissance des formes, couramment qualifié de reconnaissance des formes statistique [Fukunaga, 72], se base sur des propriétés numériques des formes étudiées. C'est dans ce deuxième contexte que se situe notre étude.

En reconnaissance des formes statistique chaque observation est représentée par un vecteur  $\mathbf{x}$  de  $d$  paramètres réels, appelé vecteur forme tel que  $\mathbf{x} = (x_1 \dots x_j \dots x_d)^t$ . Ce vecteur sera représenté par un point dans l'espace de dimension  $d$  (figure 2.1),  $\mathbb{R}^d$ , autrement appelé espace de représentation. Par exemple, dans notre application (chapitre 3), 31 paramètres ont été calculés pour définir le vecteur forme d'où une "représentation" en dimension 31.

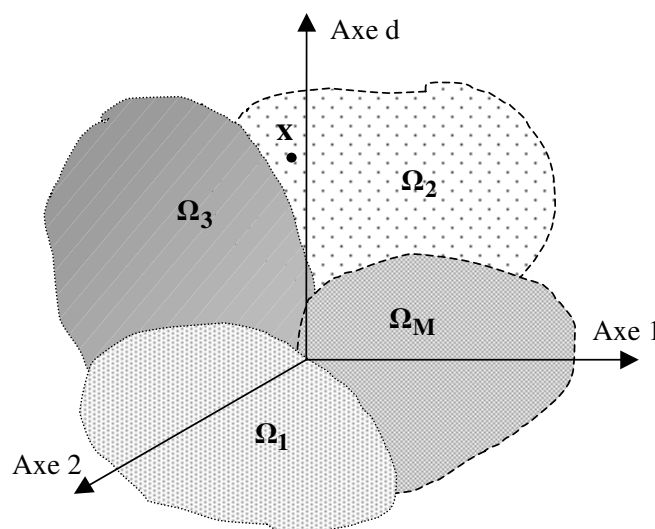
Si on se place dans le contexte du diagnostic, les paramètres du vecteur forme traduisent l'état du système étudié. Ils sont issus d'analyses effectuées sur les signaux fournis par les capteurs implantés (vibrations, vitesse, courants, ou encore tensions par exemple).

Supposons que pour toute nouvelle forme nous ayons à décider parmi  $M$  formes types, considérées comme étant des prototypes. A cause du bruit de mesure, de la précision des capteurs etc..., une nouvelle observation sera rarement identique à l'un des prototypes (points confondus dans  $\mathbb{R}^d$ ). Les classes  $(\omega_1, \omega_2, \dots, \omega_M)$  correspondent à des zones dans l'espace, regroupant les formes semblables. Ces zones de l'espace de représentation sont notées  $(\Omega_1, \Omega_2, \dots, \Omega_M)$ .

L'objectif en reconnaissance des formes est alors de décider à laquelle des classes  $\omega_1, \omega_2, \dots, \omega_M$  doit être associée une nouvelle forme. Cela revient à rechercher dans quelle zone de l'espace se situe la nouvelle observation (figure 2.2). On peut toujours faire le parallèle avec l'objectif en diagnostic qui est de définir à quel mode de fonctionnement correspond une observation recueillie sur le système.



**Fig. 2.1.** Caractérisation d'une observation par un vecteur forme représenté par un point dans  $\mathbb{R}^d$  (avec  $d = 3$ )

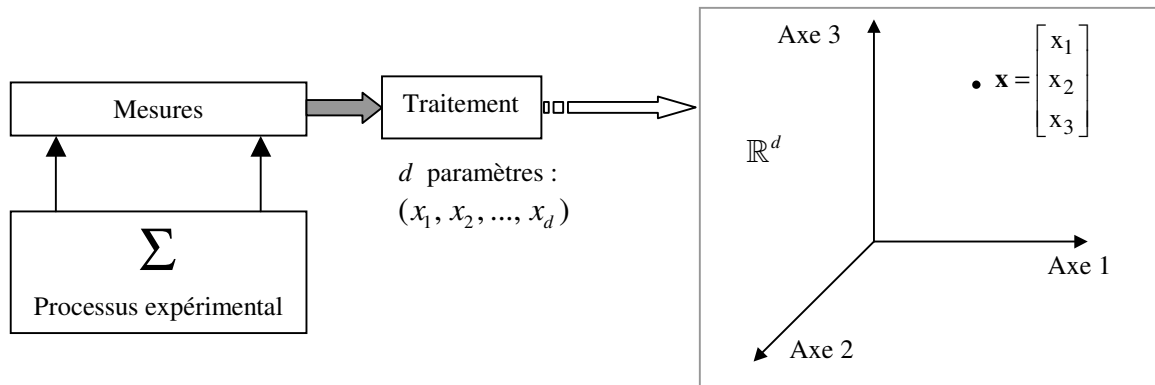


**Fig. 2.2.** Objectif en Reconnaissance des Formes : associer une nouvelle observation  $\mathbf{x}$  à l'une des classes.



L'élaboration d'un système de reconnaissance des formes se déroule en trois phases :

- Une phase d'analyse (figure 2.3) au cours de laquelle il s'agit d'étudier les informations fournies par les capteurs implantés sur le système. Si ces informations sont sous formes de signaux, alors il est nécessaire d'en extraire des valeurs (ou paramètres) numériques. Ces paramètres, qui par ailleurs constituent le vecteur forme, doivent pouvoir décrire le comportement du système.



**Fig. 2.3.** Phase d'analyse : caractérisation de l'état du système (à un instant donné) par le vecteur forme.

De cette phase d'analyse doit également sortir la définition précise des classes qui représenteront les différents modes de fonctionnement. On dispose alors d'un ensemble de  $N$  observations  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$  regroupées en  $M$  classes. C'est l'ensemble d'apprentissage. La phase d'analyse est lourde en terme de calcul et nécessite toute la connaissance du système étudié pour rechercher les paramètres adéquats par les méthodes de traitement adaptées.

- Une phase de choix du système de détection qui consiste à développer une règle de décision qui établira des frontières entre les différentes classes. La règle de décision décidera d'affecter ou non une nouvelle observation à l'une des classes connues.
- Une phase d'exploitation qui consiste à implémenter le système décisionnel qui proposera une solution pour toute nouvelle observation recueillie sur le système. La bonne exploitation du système décisionnel dépend de la pertinence du vecteur forme et des performances de la règle de décision.

Le diagnostic par reconnaissance des formes a monté son intérêt sur différentes applications industrielles [Zieba, 95], [Boutleux, 96], [Ouaghram, 99].

## 2.2. Analyse des observations

### 2.2.1. Détermination de l'espace de représentation

Il s'agit, dans cette phase de prétraitement, de construire le vecteur forme. Les  $d$  paramètres obtenus sont issus de différentes analyses effectuées sur les signaux recueillis à l'aide des capteurs. L'intervention d'un expert du processus à surveiller est souvent très utile pour aiguiller cette procédure. Ainsi chaque observation effectuée à un instant donné sera caractérisée par l'ensemble de paramètres. Les performances du système de diagnostic dépendront de la pertinence des paramètres calculés. Il est donc préférable de disposer de paramètres variant, de manière significative, en fonction des différents états du système. Les  $N$  observations (ou vecteurs)  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$  recueillies sur le système constituent l'ensemble d'apprentissage.

Le nombre souvent élevé de paramètres peut être pénalisant en terme de temps de calcul. De plus tous les paramètres calculés ne seront pas forcément pertinents vis-à-vis des modes étudiés. Des méthodes doivent donc être mises au point afin de ne conserver que les paramètres les plus représentatifs.

### 2.2.2. Réduction de l'espace de représentation

La réduction de l'espace de représentation permettra de diminuer le temps de classification pour toute nouvelle observation et d'éviter la dégradation des performances de la règle de décision. Les paramètres sélectionnés vont jouer un rôle important dans la phase d'apprentissage. En effet, ceux-ci devront caractériser au mieux les différents modes de fonctionnement. En d'autres termes, les paramètres jugés comme étant les plus pertinents seront ceux qui permettront de distinguer les classes dans l'espace de représentation. La réduction de l'espace de représentation peut être réalisée soit par des méthodes d'extraction de paramètres soit par des méthodes de sélection de paramètres.

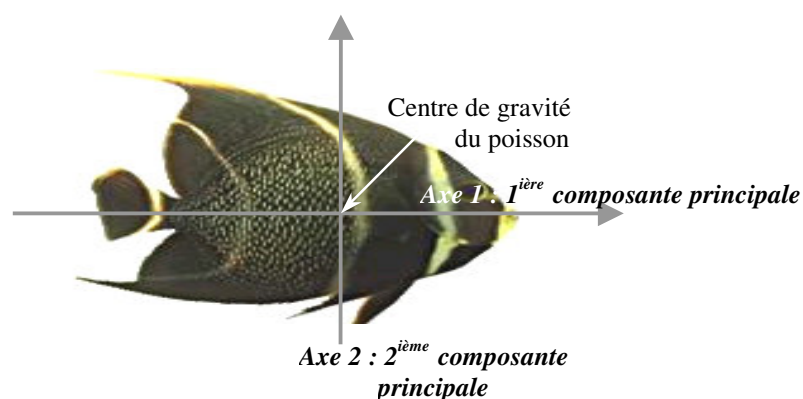
L'extraction de paramètres consiste en la recherche de  $d'$  paramètres ( $d' < d$ ), en fonction des paramètres initiaux. Les points sont alors projetés dans un sous-espace  $\mathbb{R}^{d'}$ . Le nombre de paramètres à calculer, pour caractériser chaque observation, sera toujours égal à  $d$ . Seulement, l'extraction de paramètres fournit des combinaisons linéaires de ces paramètres initiaux. Parmi les différentes méthodes d'extraction de paramètres, on distingue l'Analyse en

Composantes Principales (ACP) ou transformée de Karhunen - Loève discrète.

La sélection de paramètres consiste à rechercher les paramètres les plus pertinents et à conserver ceux-ci pour définir l'espace de représentation. Ces paramètres sélectionnés devront rendre optimum un critère de "séparabilité" entre classes et de compacité des classes comme par exemple le critère de Fisher [Fisher,36].

### 2.2.2.1 Analyse en Composantes Principales

L'analyse en composantes principales (ACP) est une technique qui permet de caractériser un système complexe multidimensionnel en dimension inférieure. Géométriquement, ce processus de mise en facteur revient à placer des axes dans un ballon de rugby. Imaginons un nuage de points en dimension 100. Dans le cas de cette dimension très élevée, il nous serait bien entendu difficile de rassembler suivant l'axe 1 (ou première composante principale) suffisamment d'information sur cet "hyper ballon". Il est donc nécessaire de définir un deuxième axe qui apporterait une bonne partie de l'information restante. Par convention, cet axe ou deuxième composante principale est perpendiculaire à la première. Si l'hyper – ballon est de forme aplatie comme "l'angélique français<sup>(1)</sup>" de la figure 2.4, alors le premier axe relierait la tête à la queue de l'animal et le deuxième axe (perpendiculaire au premier) apporterait une partie de l'information restante. Tous les axes suivants seraient perpendiculaires aux précédents et apporteraient de moins en moins d'information (sur la forme du poisson).



**Fig. 2.4.** Illustration de l'information apportée par les composantes principales.

[<http://www.ulg.ac.be/pedaexpe/cours/glosaire/acp.htm>]

<sup>(1)</sup> Angélique français

Les poissons-anges, ou angéliques, constituent, avec les poissons-papillons, la famille des chétodontidés. Ces poissons des mers tropicales, qui affectionnent les récifs coralliens, arborent souvent de superbes couleurs. L'angélique français (*Pomacanthus paru*), à la robe noire rayée de jaune, se rencontre dans les régions chaudes de l'océan Atlantique, notamment aux Antilles. *Encyclopédie Encarta Dorling Kindersley*



Dans le cadre de notre étude, le but est de définir un sous-espace de dimension  $d'$ , inférieure à  $d$ , pour mieux représenter l'ensemble d'apprentissage [Govaert, 03].

Soit  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$  l'ensemble des  $N$  vecteurs (matrice des données) de dimension  $d$ . Si les composantes de ces vecteurs ne sont pas homogènes, alors les résultats de l'ACP et de la classification risquent d'être biaisés.

Afin de parer à cette éventualité les données sont normalisées de la manière suivante :

Le centre de gravité de l'ensemble d'apprentissage est le vecteur défini par :

$$\mathbf{m} = \frac{1}{N} \sum_{k=1}^N \mathbf{x}_k \quad (2.1)$$

et les composantes de la variance sont :

$$\sigma_j^2 = \frac{1}{N} \cdot \sum_{k=1}^N (x_{kj} - m_j)^2 \quad (2.2)$$

avec  $j=1 \dots d$ .

Ainsi, à partir des deux relations précédentes on en déduit les vecteurs formes normalisés (c'est-à-dire de moyenne nulle et d'écart type unité) :

$$\mathbf{x}_{k \text{ norm}} = \Delta^{-1/2} (\mathbf{x}_k - \mathbf{m}) \quad (2.3)$$

Avec  $\Delta$  : la matrice diagonale des variances  $\sigma_j^2$  :

$$\Delta = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_d^2 \end{bmatrix} \quad (2.4)$$

Pour la suite de l'étude, les données seront considérées comme étant normalisées. La base du sous-espace est définie par les vecteurs propres de la matrice de variance – covariance [Pinkowski, 97]:

$$\mathbf{S} = \frac{1}{N} \sum_{k=1}^N \mathbf{x}_k \cdot \mathbf{x}_k^t \quad (2.5)$$

$\mathbf{S}$  une matrice carrée ( $d \times d$ ), réelle et symétrique, c'est aussi la matrice d'inertie du nuage de points considéré. Ses éléments diagonaux sont les variances de l'ensemble d'apprentissage.

Elles valent 1 car les données sont centrées et réduites. Et les éléments non diagonaux sont les covariances.

Les vecteurs propres (2.6) de la matrice de variance- covariance  $\mathbf{S}$  sont les vecteurs  $\mathbf{v}_j$  tels que  $\mathbf{S} \cdot \mathbf{v}_j = \lambda_j \cdot \mathbf{v}_j$ . Les scalaires  $\lambda_j$ , repris dans la matrice diagonale (2.7), sont les valeurs propres.

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_{d'} & \dots & \mathbf{v}_d \end{bmatrix} \quad (2.6)$$

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 & & & & & 0 \\ & \lambda_2 & & & & \\ & & \dots & & & \\ & & & \lambda_{d'} & & \\ & & & & \dots & \\ 0 & & & & & \lambda_d \end{bmatrix} \quad (2.7)$$

Les premiers  $d'$  vecteurs propres, rangés par ordre décroissant des valeurs propres vont constituer la base du sous-espace de dimension  $d'$ . Le vecteur  $\mathbf{x}_j^P$  dont les  $N$  composantes sont les projections des  $N$  vecteurs  $\mathbf{x}_k$  sur le vecteur  $\mathbf{v}_j$  est appelé  $j^{\text{ième}}$  composante principale.

Chaque composante  $x_{kj}^P$  du vecteur  $\mathbf{x}_j^P$  aura pour expression :

$$x_{kj}^P = \mathbf{x}_k^t \cdot \mathbf{v}_j \text{ avec } k = 1, \dots, N \quad (2.8)$$

La nouvelle matrice des données s'écrit :

$$\mathbf{X}^P = \begin{bmatrix} \mathbf{x}_1^P & \mathbf{x}_2^P & \dots & \mathbf{x}_{d'}^P & \dots & \mathbf{x}_d^P \end{bmatrix} = \mathbf{X}^t \mathbf{V} \quad (2.9)$$

Les axes définis par ces vecteurs  $\mathbf{x}_j^P$  sont dits factoriels, et sont les vecteurs propres. Ces axes sont les combinaisons linéaires des  $d$  axes initiaux.

On aboutit ainsi à une représentation optimale de la matrice des données  $\mathbf{X}$  dans un espace de dimension inférieure. L'ACP garantit que la déformation du nuage de points est minimale au sens d'un critère  $J_p$  [Dubuisson, 90] basé sur la moyenne des carrés des distances entre les

vecteurs  $\mathbf{x}_k$  et leurs projections notées  $\mathbf{x}_k^P$  (telles que  $\mathbf{x}_k^P = [x_{k1}^P, x_{k2}^P, \dots, x_{kd}^P]^t$ ) :

$$J_p = \frac{1}{N} \sum_{k=1}^N d^2(\mathbf{x}_k, \mathbf{x}_k^p) \quad (2.10)$$

Les résultats de l'ACP peuvent être visualisés en considérant la projection du nuage de points sur les deux ou trois premiers axes factoriels. Autrement dit, en visualisant la matrice des données dans l'espace défini par les deux ou trois premiers vecteurs propres.

Enfin la qualité de la représentation peut être quantifiée par le calcul du pourcentage d'inertie expliqué sur chaque axe factoriel ou sur le sous-espace de dimension  $d'$  obtenu. Ce pourcentage d'inertie est défini en fonction des valeurs propres :

- Pourcentage d'inertie sur l'axe  $i$  :

$$\Gamma_i = \frac{\lambda_i}{\sum_{j=1}^d \lambda_j} \quad (2.11)$$

- Qualité de la représentation dans le sous-espace de dimension  $d'$  :

$$\Gamma = \frac{\sum_{i=1}^{d'} \lambda_i}{\sum_{j=1}^d \lambda_j} \quad (2.12)$$

Ce rapport doit tendre vers 1 pour une représentation la plus proche possible de la réalité. Mais en pratique l'utilisateur fixe un seuil de qualité autour de 80%, 90%, puis sélectionne les vecteurs propres correspondant aux valeurs propres ayant permis d'atteindre ce seuil, afin de définir la base du sous-espace.

Une fois le sous-espace  $\mathbb{R}^{d'}$  défini, il est possible d'évaluer la qualité de la représentativité de la projection  $\mathbf{x}_k^p$  du vecteur  $\mathbf{x}_k$ , dans ce sous-espace, en calculant le cosinus de l'angle  $\alpha_k$  que forme ce dernier avec  $\mathbb{R}^{d'}$  :

$$\cos^2(\alpha_k) = \frac{\sum_{i=1}^{d'} (x_{ki}^p)^2}{\sum_{j=1}^d (x_{kj})^2} \quad (2.13)$$

Ainsi, plus  $\cos^2(\alpha_k)$  est proche de l'unité plus la projection de  $\mathbf{x}_k$  en  $\mathbf{x}_k^p$  est

représentative.

Des résultats d'ACP peuvent être consultés dans le chapitre 3.

### 2.2.2.2 Sélection de paramètres

L'ACP ou plus généralement les méthodes d'extraction de paramètres permettent de construire un sous-espace de dimension inférieure à partir de l'espace initial pour entre autres décrire et visualiser l'ensemble d'apprentissage. Seulement cette procédure ne permet pas de réduire le nombre de paramètres à calculer pour chaque observation (étant donnée que les composantes des vecteurs projetés sont des combinaisons linéaires des composantes des vecteurs initiaux). Or réduire le nombre de paramètres en ne sélectionnant que les plus discriminants offre plusieurs avantages :

- le temps de calcul est réduit non seulement pour l'extraction de paramètres, mais aussi pour la décision concernant toute nouvelle observation.
- les performances de la classification sont améliorées.

Par la suite l'utilisation de l'analyse en composantes principales sera essentiellement réservée à la visualisation des données de l'ensemble d'apprentissage (chapitre 3).

Pour réduire le nombre de paramètres et, par la même occasion, la dimension de l'espace de représentation on utilise des méthodes de sélection de paramètres.

L'objectif est de rechercher, parmi les  $d$  paramètres de départ, un ensemble de  $d'$  paramètres donnant le plus d'information. Toutefois, la sélection nécessite la connaissance de la structure en classes des données pour être utilisée de la manière la plus efficace possible. En d'autres termes, l'utilisateur doit connaître le nombre de classes existantes ainsi que les observations correspondantes.

En outre le sous-ensemble de paramètres devra maximiser un critère rendant compte à la fois :

- d'une faible dispersion des points dans chaque classe, ce qui suggère une variance intra-classe minimale.
- d'un éloignement maximal entre les différentes classes, donc d'une variance inter – classes maximale.

Les matrices de variance intra – classe et inter – classes traduisent ces deux conditions. En effet l'expression de la matrice de variance intra-classe fait intervenir les échantillons de chaque classe ainsi que les centres de gravités :



$$\mathbf{S}_W = \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^{N_i} (\mathbf{x}_{ij} - \mathbf{m}_i)(\mathbf{x}_{ij} - \mathbf{m}_i)^t \quad (2.14)$$

Alors que la matrice de dispersion inter-classes s'exprime en fonction des centres de gravités des classes et du centre de gravité de l'ensemble d'apprentissage :

$$\mathbf{S}_B = \frac{1}{N} \sum_{i=1}^M (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^t \quad (2.15)$$

Avec :

$M$  : Nombre de classes

$N$  : Nombre de vecteurs de l'ensemble d'apprentissage

$N_i$  : Nombre de vecteurs dans la classe  $\omega_i$

$\mathbf{m}_i$  : Centre de gravité de la classe  $\omega_i$

$\mathbf{m}$  : Centre de gravité de l'ensemble d'apprentissage

$\mathbf{x}_{ij}$  :  $j^{\text{ième}}$  vecteur de la classe  $\omega_i$

Par ailleurs, la matrice variance-covariance qui traduit la dispersion totale de l'ensemble vaut:

$$\mathbf{S} = \mathbf{S}_W + \mathbf{S}_B \quad (2.16)$$

Le critère de Fisher appliqué à chacun des paramètres permet de quantifier le pouvoir de discrimination de ceux-ci. Dans le cas de deux classes et pour un paramètre  $\alpha$  donné, le critère s'exprime de la manière suivante :

$$J(\alpha) = \frac{\mathbf{m}_1(\alpha) - \mathbf{m}_2(\alpha)}{N_1 \sigma_1^2(\alpha) - N_2 \sigma_2^2(\alpha)} \quad (2.17)$$

Avec :

- $\mathbf{m}_i(\alpha)$  ( $i = 1, 2$ ), centre de gravité de la classe  $\omega_i$  en ne considérant que le paramètre  $\alpha$
- $\sigma_i^2(\alpha)$  : variance de la composante  $\alpha$  des vecteurs de la classe  $\omega_i$  :

$$\mathbf{m}_i(\alpha) = \frac{1}{N_i} \sum_{k_i=1}^{N_i} \mathbf{x}_{k_i}(\alpha) \quad (2.18)$$

$$\sigma_i^2(\alpha) = \frac{1}{N_i} \sum_{j=1}^{N_i} [\mathbf{x}_{ij}(\alpha) - \mathbf{m}_i(\alpha)]^2 \quad (2.19)$$

D'une manière générale, pour M classes,  $J(\alpha)$  s'écrit :

$$J(\alpha) = \sum_{i=1}^M \sum_{j=1}^{M-1} \frac{\mathbf{m}_i(\alpha) - \mathbf{m}_j(\alpha)}{N_i \sigma_i^2(\alpha) - N_j \sigma_j^2(\alpha)} \quad (2.20)$$

Le numérateur de l'expression (2. 20) traduit la séparation des classes (variance inter – classes) alors que le dénominateur reflète la compacité des classes (par analogie avec la variance intra – classe). Plus le critère est grand, plus les classes sont séparées.

Ainsi une première solution pour la sélection des paramètres consiste à prendre en compte les  $d'$  premiers paramètres vérifiant :

$$J(\alpha_1) \geq J(\alpha_2) \geq \dots \geq J(\alpha_{d'}) \geq J(\alpha_d) \quad (2.21)$$

Pour définir la dimension finale  $d'$  il suffit par exemple de visualiser la valeur du critère pour chacun des paramètres. La figure suivante donne un exemple de calcul du critère. Cette figure nous permet de dire que deux paramètres ( $\alpha_1$  et  $\alpha_3$ ) concourent activement à la compacité et à la séparabilité des classes. Un troisième paramètre ( $\alpha_2$ ) intervient dans une moindre mesure, suivi de quatre autres paramètres ( $\alpha_{31}, \alpha_{29}, \alpha_{28}$  et  $\alpha_{25}$ ) qui ont une influence beaucoup plus faible. Les autres paramètres ne sont pas intéressants dans ce cas de figure.

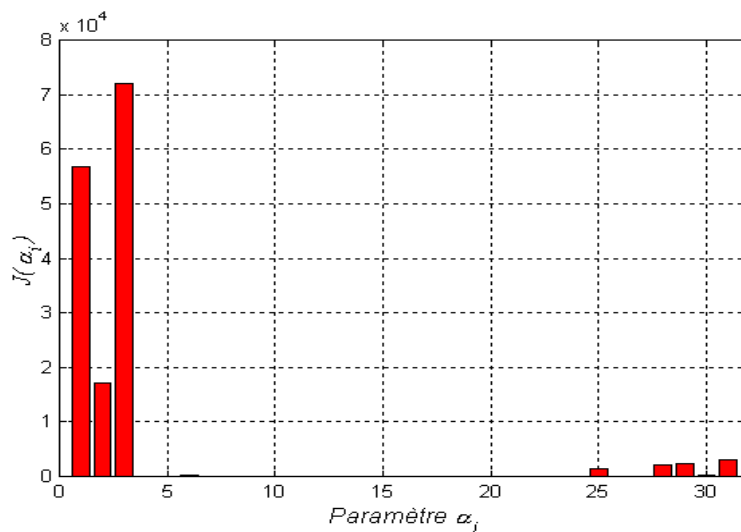


Fig. 2.5. Exemple d'analyse permettant de choisir la dimension du vecteur forme.

Nous reviendrons plus en détail sur ce calcul et sur l'application directe correspondante dans le prochain chapitre.

Mais bien qu'étant particulièrement intéressante pour mettre en avant la pertinence des paramètres, cette méthode n'est pas forcément optimale quant à la recherche d'un sous-ensemble  $E'$  (de  $d'$  paramètres) pour satisfaire au mieux les conditions de séparabilité et de compacité des classes. En effet, elle ne montre que l'apport individuel des paramètres.

Si on considère les matrices de dispersion intra – classe et inter – classes (2. 14) et (2. 15), alors deux autres critères rendant compte directement de la pertinence des sous-ensembles de paramètres peuvent être utilisés :

$$J' = \text{trace}(\mathbf{S}_W^{-1} \mathbf{S}_B) \text{ et } J'' = \frac{\det(\mathbf{S})}{\det(\mathbf{S}_W)} \quad (2. 22)$$

Ainsi, il est possible de rechercher un ensemble  $E'$  de  $d'$  paramètres pertinents de telle sorte que :

$$J'(E') = \text{Max}_{i=1, \dots, n_c} \left[ J(E_{d'}^{(i)}) \right] \quad (2. 23)$$

Avec  $n_c$  : le nombre de combinaisons (de  $d'$  paramètres parmi  $d$ ) étudiées, et  $E_{d'}^{(i)}$  la  $i^{\text{ème}}$  combinaison de  $d'$  paramètres pour laquelle on calcul la valeur du critère.

La première solution serait de calculer la valeur du critère pour toutes les combinaisons possibles de  $d'$  paramètres. Mais le nombre total de combinaisons est égale à  $C_d^{d'}$ , ce qui devient vite dissuasif si la dimension de départ est élevée.

D'autant plus qu'il existe des méthodes sous-optimales qui, même si elles ne permettent pas une recherche exhaustive, fournissent des résultats réalistes. On peut ici en citer quelques-unes [Kudo, 00] :

- La méthode appelée SFS ("Sequential Forward Selection") effectue une sélection des paramètres de manière successive. A chaque étape on retient le paramètre qui, s'il est pris en compte avec l'ensemble des paramètres déjà sélectionnés à l'étape précédente, réalise une optimisation du critère. L'ensemble de départ est donc l'ensemble vide.
- La méthode GSFS ("Generalised Sequential Forward Selection") fonctionne comme l'algorithme précédent à ceci près qu'on sélectionne plusieurs paramètres à chaque itération.

- La méthode SBS ("Sequential Backward Selection") est celle qui a été utilisée dans le cadre des travaux. Contrairement à la méthode SFS, le sous-espace de départ est l'ensemble des paramètres. D'une manière plus détaillée, la procédure est la suivante :

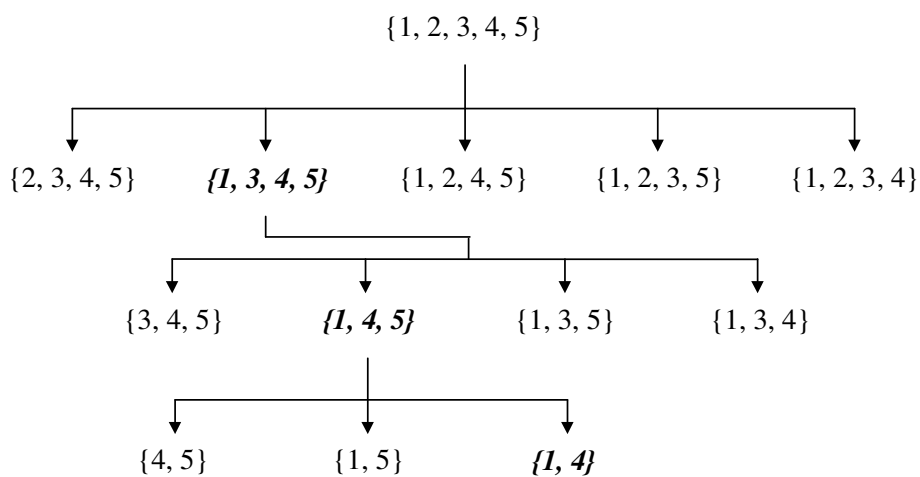
A l'étape k on calcule le critère pour les sous-ensembles  $E_{k,i}$  de dimension  $d - k$ . Les sous-espaces  $E_{k,i}$  sont issus du sous-espace  $E_{k-1}$  en omettant volontairement le paramètre  $\alpha_i$  :

$$E_{k-1} \supset E_{k,i} = \{E_{k-1} - \alpha_i\} \quad (2.24)$$

Ainsi le sous-espace  $E_k$  pris en considération pour l'étape suivante est celui pour lequel le critère est optimal :

$$J(E_k) = \text{Max}_{i=1, d-k} [J(E_{k,i})] \quad (2.25)$$

La figure 2.6 illustre cette méthode avec  $d = 5$  et  $d' = 2$ . A chaque étape on sélectionne le sous-ensemble de paramètres pour lequel le critère est maximal. Les numéros représentent la position des paramètres dans le vecteur forme de départ. Les paramètres sélectionnés sont  $\alpha_1$  et  $\alpha_4$ .



**Fig. 2.6.** Schéma de principe de la sélection de paramètres par la méthode SBS, avec  $d=5$  et  $d'=2$ .

Le caractère non optimal de cette méthode vient, entre autres, du fait que le critère n'a pas été calculé pour toutes les combinaisons de  $d'$  paramètres et que par conséquent la sélection finale obtenue n'est pas forcément la meilleure. De plus, cette procédure nécessite, à chaque étape, le calcul du critère pour toutes les combinaisons prises en considération. Avec cette

méthode, le nombre de combinaisons vaut:

$$\sum_{j=1}^{d-d'} C_{d-j}^1 \quad (2.26)$$

Ce qui représente tout de même un gain de temps par rapport à une recherche exhaustive.

- Enfin la méthode GSBS ("Generalised Sequential Backward Selection") utilise la même procédure que la méthode SBS mais en omettant plus d'un paramètre, pour obtenir les sous-espaces, à chaque étape.

Il existe des méthodes de sélection optimales telles les algorithmes basés sur une optimisation de type *Branch and Bound*. Ces méthodes prennent implicitement en compte toutes les combinaisons de paramètres en utilisant un arbre de sélection dont le parcours n'est pas nécessairement complet. En effet, en vertu de l'hypothèse de la monotonie du critère  $J$  (2.21), la valeur obtenue pour un sous-ensemble de  $d'$  paramètres peut ne pas être remise en cause. Cela entraîne la non prise en compte des solutions que pourraient fournir les autres combinaisons et donc un gain de temps [Somol, 00], [Chen, 03].

### 2.3. Détermination de l'espace de décision

Il s'agit, au cours de cette étape, de rechercher une structure en classes de l'ensemble des données par la mise en place de procédures de classification. La classification peut être soit exclusive soit non exclusive. Dans le premier cas une observation appartient à un seul sous-ensemble, alors que dans le second cas une même observation peut appartenir à plusieurs sous-ensembles. L'algorithme fournit alors une valeur numérique de l'appartenance de l'observation à chaque sous-ensemble.

Lorsqu'on connaît la classe d'origine de chaque observation l'espace de décision est parfaitement connu et l'apprentissage peut être effectué en mode dit supervisé [Nong, 01]. En terme de diagnostic, une telle approche permettrait de vérifier qu'on retrouve bien les différentes classes ou modes de fonctionnement.

Si au contraire on ne dispose d'aucune information sur la structuration de l'ensemble d'apprentissage en classes, alors la classification doit être faite en mode non supervisé. Il existe plusieurs méthodes de classification non supervisée, autrement appelées méthodes de coalescence.

Si on se limite à l'affectation exclusive ces méthodes se décomposent en deux parties :

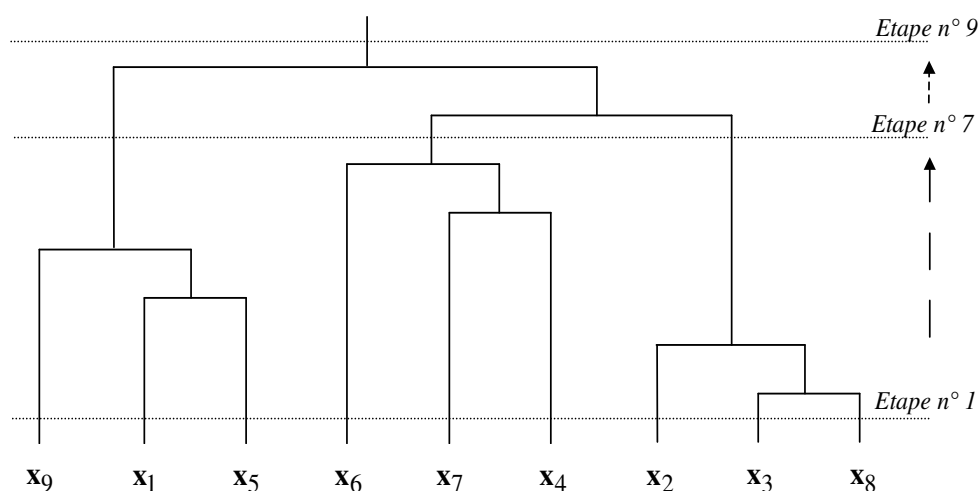
- les méthodes hiérarchiques (ascendantes ou descendantes)

- les méthodes de classification par partition

### 2.3.1. Méthodes de classification hiérarchiques

La classification hiérarchique utilise les distances entre observations ou ensembles d'observations lors de la formation des classes. On obtient une arborescence avec à la base les observations elles mêmes et, au sommet l'ensemble des données regroupées en une seule classe. Cette arborescence contient autant de niveaux que d'observation de l'ensemble d'apprentissage. L'algorithme de classification ascendante procède à un regroupement progressif des éléments, alors que la classification descendante vise plutôt l'éclatement des sous-ensembles.

Pour illustrer ces méthodes, la figure 2.7 présente le schéma de principe d'une coalescence hiérarchique ascendante. L'ensemble d'apprentissage est constitué au départ (étape n°1) des 9 échantillons, donc de 9 "classes".



**Fig. 2.7.** Schéma de principe de la classification hiérarchique ascendante. L'ensemble d'apprentissage est constitué de 9 observations.

Dans cet exemple, la procédure peut être arrêtée à un niveau approprié tel que l'étape n° 7. On obtient alors un regroupement en trois classes constituées des observations suivantes :  $\{x_2, x_3, x_8\}$ ,  $\{x_4, x_6, x_7\}$  et  $\{x_1, x_5, x_9\}$ . L'algorithme peut être contrôlé (arrêt) par un critère de compacité – séparabilité des classes ou validé après visualisation.

Ces algorithmes sont évidemment coûteux en temps de calcul et en espace mémoire pour un nombre élevé d'échantillons. Toutefois de récents travaux [Moghaddam, 03] [Sonbaty, 98] y ont apporté des améliorations afin que ces algorithmes soient plus rapides et utilisables en ligne.

### 2.3.2. Méthodes de classification par partition

Ces méthodes sont adaptées à la définition selon laquelle les classes sont "des zones de l'espace regroupant les formes semblables". En effet, les méthodes de coalescence réalisant une classification par partition ont pour but de créer des regroupements d'observations autour de centres de gravité recherchés. Les algorithmes utilisés font partie de la famille dite des méthodes des nuées dynamiques introduites par [Diday, 76]. Parmi ces algorithmes, on distingue tout particulièrement la coalescence stricte Hard C-Means (HCM) introduite par [Duda, 73]. Comme son nom l'indique, cet algorithme produit des frontières strictes entre les classes considérées.

L'algorithme HCM fonctionne de la manière suivante :

Initialisation :

- Choix du nombre de classes  $M$
- Initialisation aléatoire de la matrice de partition  $U$

La matrice de partition  $U$  à actualiser à chaque itération est composée des coefficients d'appartenance  $\mu_{ki}$  telle que :

$$U = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1i} & \cdots & \mu_{1M} \\ \vdots & & \vdots & & \vdots \\ \mu_{k1} & \cdots & \mu_{ki} & \cdots & \mu_{kM} \\ \vdots & & \vdots & & \vdots \\ \mu_{N1} & \cdots & \mu_{Ni} & \cdots & \mu_{NM} \end{bmatrix} \quad (2.27)$$

Avec :

$$\mu_{ki} \in \{0,1\} \text{ tel que :} \quad (2.28)$$

$$\forall i, \mu_{ki} = 1 \text{ et } \forall j \neq i, \mu_{kj} = 0$$

A chaque itération  $q$  :

- Calcul des centres de gravité  $\mathbf{m}_i$  de chacune des classes tels que :

$$\mathbf{m}_i = \frac{\sum_{k=1}^N \mu_{ki} \mathbf{x}_k}{\sum_{k=1}^N \mu_{ki}} \quad (2.29)$$

- Calcul des distances entre les observations et les centres de gravité  $\mathbf{m}_i$  :

$$d(\mathbf{x}_k, \mathbf{m}_i) = [\mathbf{x}_k - \mathbf{m}_i]^t [\mathbf{x}_k - \mathbf{m}_i] = \left[ \sum_{j=1}^d (x_{kj} - m_{ij})^2 \right]^2 \quad (2.30)$$

- Affectation des observations aux classes (centre de gravité le plus proche) et actualisation de la matrice  $U$  :

$$\mu_{ki}^{(q+1)} = \begin{cases} 1 & \text{si } d(\mathbf{x}_k, \mathbf{m}_i) = \min_{j=1, M} [d(\mathbf{x}_k, \mathbf{m}_j)] \\ 0 & \text{sinon} \end{cases} \quad (2.31)$$

- Calcul de l'indice d'évolution :  $\mathcal{E} = \|U^{(q+1)} - U^{(q)}\|$ .

Tant que  $\mathcal{E} > 0$  alors les données (2.28) à (2.30) sont calculées pour l'itération suivante. Si  $\mathcal{E}$  devient nul alors la partition est restée inchangée entre les itérations (q) et (q+1). En d'autres termes, cela signifie que l'algorithme de coalescence a défini une structure en classes de l'ensemble et que celle-ci n'évoluera plus. D'une manière intrinsèque, à la fin de la coalescence la variance intra-classe est minimisée. On peut la quantifier grâce au calcul du critère  $J_w$  qui correspond à la trace de la matrice de dispersion intra-classe :

$$J_w = \sum_{k=1}^N \sum_{i=1}^M \mu_{ki} \cdot d(\mathbf{x}_k, \mathbf{m}_i)^2 \quad (2.32)$$

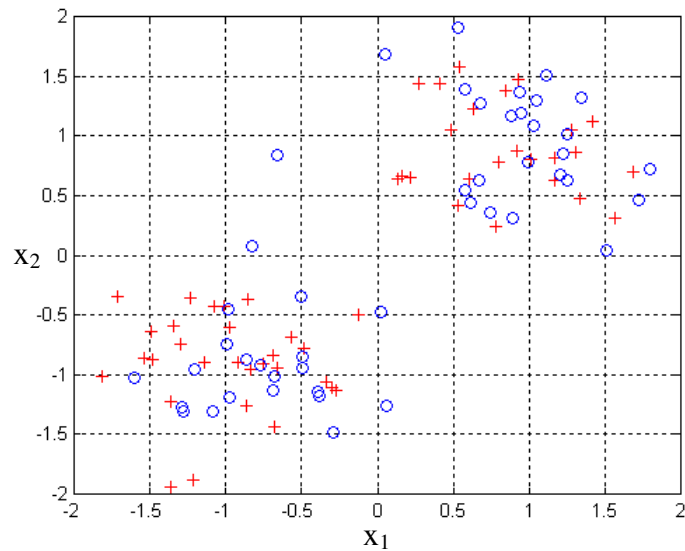
Cet algorithme est illustré à travers l'exemple suivant :

Soit un ensemble de 100 données artificielles de dimension 3 dont la partition aléatoire de départ est présentée sur la figure 2.8 et, le résultat de la coalescence sur la figure 2.9. En réalité, l'ensemble est constitué de deux classes de 50 échantillons chacune, dont les caractéristiques (moyennes et écart-types) sont les suivantes :

$$\text{Classe } \omega_1 : \begin{cases} N_1 = 50 \\ \mathbf{m}_1 = [-0.88 \quad -0.91 \quad -0.9] \\ \boldsymbol{\sigma}_1 = [0.45 \quad 0.39 \quad 0.44] \end{cases}$$

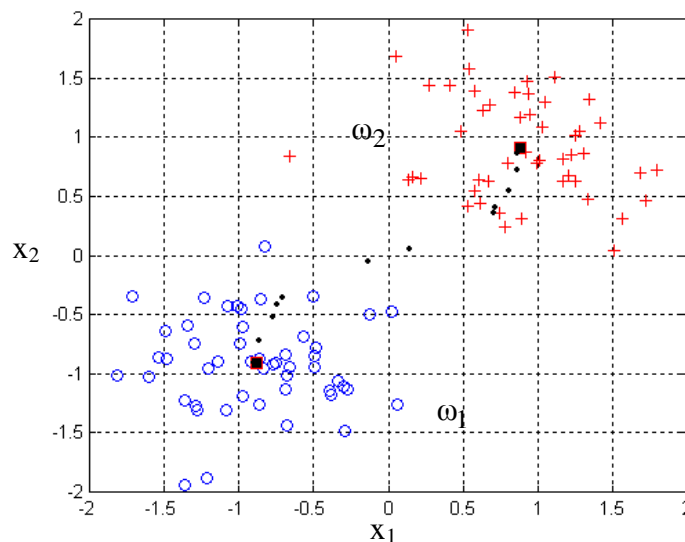
$$\text{Classe } \omega_2 : \begin{cases} N_2 = 50 \\ \mathbf{m}_2 = [0.88 \quad 0.91 \quad 0.9] \\ \boldsymbol{\sigma}_2 = [0.47 \quad 0.42 \quad 0.39] \end{cases}$$





**Fig. 2.8.** Résultat de la partition aléatoire de l'ensemble. 46 échantillons ont été affectés à la classe  $\omega_1$  ("0") et 54 à la classe  $\omega_2$  ("+").

A l'issue de la phase d'initialisation (partition aléatoire) l'ensemble d'apprentissage a été séparé en deux classes de 46 et 54 échantillons. L'algorithme a convergé en 7 itérations (figure 2.9) et la structure en classe a bien été trouvée. Cette figure fournit également un historique des trajectoires (en pointillés) suivies par les deux centres de gravité.

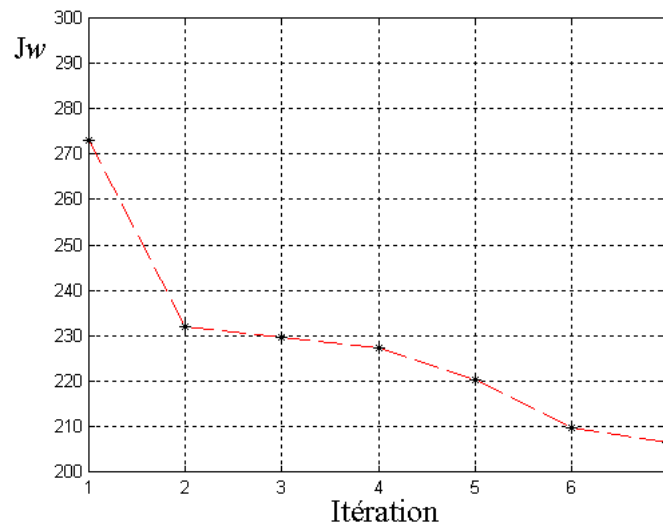


**Fig. 2.9.** Résultat de la coalescence HCM. Après 7 itérations : partition stricte en deux classes  $\omega_1$  ("0") et  $\omega_2$  ("+") de 50 échantillons et centres de gravité obtenus(■)

Il convient toutefois de préciser que l'algorithme a convergé de manière satisfaisante vers le résultat escompté parce que les deux classes sont distinctes et que leurs dispersions

respectives sont quasi identiques.

Enfin la figure 2.10 donne les différentes valeurs prise par le critère de coalescence  $J_w$  (équation (2. 32)). La décroissance du critère corrobore le fait que, à partir d'une partition aléatoire initiale, l'algorithme recherche la meilleure partition au sens de la variance intra-classe.



**Fig. 2.10.** Variation de la trace de la matrice de dispersion intra-classe au cours de la coalescence

Le résultat final de la coalescence peut dépendre de la partition aléatoire initiale. Il est donc conseillé de lancer la recherche à plusieurs reprises à partir de partitions initiales différentes. A l'issue de chaque coalescence, certaines observations risquent d'être systématiquement regroupées au sein des même classes. Ces observations, considérées comme étant solidaires, seront appelées formes fortes.

Dans le cas de classes à distributions non gaussiennes, l'algorithme n'est pas adapté. Il se peut que l'algorithme échoue, surtout si les classes sont de dispersions différentes. En effet, l'algorithme recherche des classes de formes hypersphériques car c'est la distance euclidienne qui est utilisée. Si les classes étaient gaussiennes (ou ovales tout simplement), le résultat serait moins bon.

Les différentes étapes de l'apprentissage précédemment citées peuvent être combinées pour rechercher d'une part la meilleure représentation possible des observations (le vecteur forme) et, d'autre part, la meilleure structure de l'ensemble d'apprentissage en classes.

La dernière étape, décrite au paragraphe suivant, consiste à définir une règle de décision qui permettra de classer ou non toute nouvelle observation recueillie sur le processus.

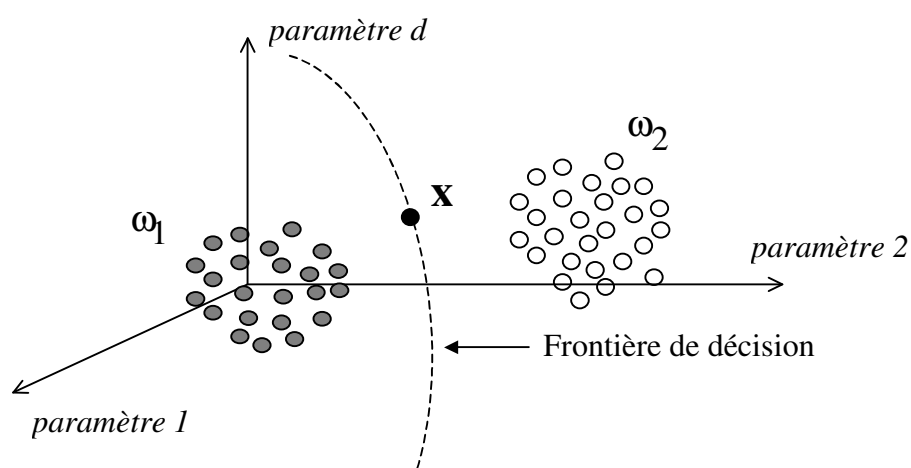
## 2.4. Choix de la procédure de décision

A l'issue des deux premières étapes l'ensemble d'apprentissage a été structuré en zones distinctes de l'espace. Les observations qui y sont regroupées sont parfaitement connues grâce à la définition du vecteur forme. Les données sont étiquetées de telle sorte que chacune des observations appartient à une des  $M$  classes connues :  $\forall_{i=1,N} \mathbf{x}_i, \exists_{j=1,M} \omega_j$  telle que  $x_i \in \omega_j$ .

Soit  $\mathbf{X}=\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$  l'ensemble d'apprentissage et,  $\{\omega_1, \omega_2, \dots, \omega_M\}$  l'ensemble des  $M$  classes (ou modes de fonctionnement en diagnostic) connues. Il s'agit maintenant de définir à quelle classe affecter une nouvelle observation  $\mathbf{x}_u$  recueillie à un instant donné sur le système.

Toutefois, l'affectation ne doit pas être systématique car des erreurs peuvent être commises. L'algorithme de décision doit donc bénéficier de solutions alternatives afin de diminuer le risque de mauvaise classification. Ces solutions sont fournies par les options de rejets d'ambiguïté et de distance. Il s'agit en réalité de différer la décision car les observations dont l'appartenance aux classes est ambiguë ou insuffisante, pourraient être précurseurs de l'apparition de nouvelles classes.

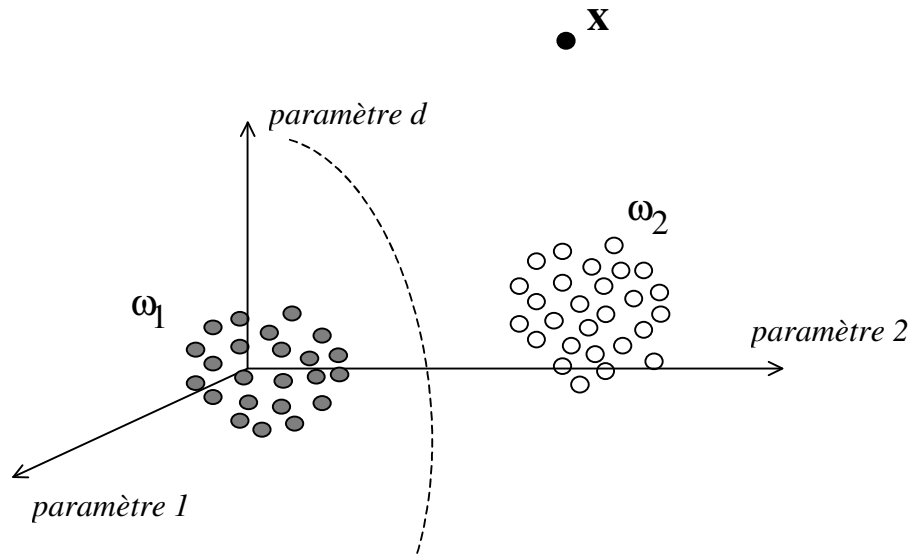
Le rejet d'ambiguïté est appliqué aux observations se situant à l'intersection de classes ou près d'une frontière entre deux classes comme le montre la figure 2.11.



**Fig. 2.11.** Observation située près de la frontière de décision entre deux classes

Les vecteurs se trouvant dans cette situation sont affectés à une nouvelle classe fictive  $\omega_0$  appelée classe de rejet d'ambiguïté.

Le rejet de distance concerne les vecteurs situés dans des zones de l'espace qui ne correspondent à aucune des classes connues de l'ensemble d'apprentissage (figure 2.12).



**Fig. 2.12.** Application du rejet de distance. La nouvelle observation  $\mathbf{x}$  ne correspond à aucune des classes connues.

Cette autre alternative au risque de mauvaise classification est nécessaire en diagnostic. Ces nouvelles mesures peuvent laisser présager l'apparition de certains modes de fonctionnement qui étaient jusque là inconnus. Les vecteurs rejetés en distance seront affectés à une nouvelle classe fictive  $\omega_d$ , appelée classe de rejet en distance.

Globalement, la règle de décision incluant les deux options de rejets sera appliquée pour  $M + 2$  classes :

$$\begin{aligned}
 \mathbf{x} &\rightarrow \omega_i \quad (i = 1, M) : \mathbf{x} \text{ est classé dans } \omega_i \\
 \mathbf{x} &\text{ est rejeté en ambiguïté alors : } \mathbf{x} \rightarrow \omega_0 \\
 \mathbf{x} &\text{ est rejeté en distance alors : } \mathbf{x} \rightarrow \omega_d
 \end{aligned}
 \tag{2.33}$$

On distingue deux types d'approche pour la mise au point d'une règle de décision. Le premier type d'approche consiste à utiliser des méthodes statistiques, paramétriques ou non paramétriques. La deuxième approche, qui est analytique, consiste à privilégier le calcul des frontières de décision entre classes.

### 2.4.1. Méthodes paramétriques

Les méthodes paramétriques supposent la connaissance des lois de probabilité des observations et des classes. En effet, tout vecteur  $\mathbf{x}$  de  $\mathbb{R}^d$  suit, dans une classe donnée  $\omega_i$ , une loi de probabilité  $f(\mathbf{x}|\omega_i)$ . Par ailleurs, les classes  $\omega_j$  ( $j=1,\dots,M$ ) sont caractérisées par leurs probabilités a priori  $\Pr(\omega_j)$ . Ces probabilités vérifient la condition d'orthogonalité suivante :

$$\sum_{j=1}^M \Pr(\omega_j) = 1 \quad (2.34)$$

Dans ce contexte, la règle de décision la plus couramment utilisée est la règle de Bayes [Dubuisson, 90].

#### 2.4.1.1 Règle de Bayes

Cette règle utilise la valeur de la probabilité a posteriori qu'une observation  $\mathbf{x}$  provienne de la classe  $\omega_i$  :

$$\Pr(\omega_i|\mathbf{x}) = \frac{f(\mathbf{x}|\omega_i) \cdot \Pr(\omega_i)}{\sum_{j=1}^M f(\mathbf{x}|\omega_j) \cdot \Pr(\omega_j)} \quad (2.35)$$

Où  $f(\mathbf{x}|\omega_i)$  est la densité de probabilité de  $\mathbf{x}$  dans la classe  $\omega_i$ .

La quantité au dénominateur est appelée densité de probabilité non conditionnelle à  $\mathbf{x}$ , ou loi de mélange. Par la suite elle sera notée  $f(\mathbf{x})$ .

La règle de Bayes définit l'affectation du vecteur  $\mathbf{x}$  à l'une des classes par :

$$\mathbf{x} \rightarrow \omega_i \text{ si } \Pr(\omega_i|\mathbf{x}) = \max_{j=1,M} [\Pr(\omega_j|\mathbf{x})] \quad (2.36)$$

#### 2.4.1.2 Rejet d'ambiguïté

Cette option de rejet peut être incluse dans la règle de Bayes par l'introduction d'un coût de rejet  $C_r$  constant, relatif au rejet d'un vecteur  $\mathbf{x}$  dans la classe  $\omega_j$  ( $j=1\dots M$ ). La règle de décision incluant le rejet en ambiguïté est définie à partir des probabilités a posteriori par :

$$\begin{cases} \mathbf{x} \rightarrow \omega_i \text{ si } \Pr(\omega_i | \mathbf{x}) = \max_{j=1, M} [\Pr(\omega_j | \mathbf{x})] \geq 1 - C_r \\ \mathbf{x} \rightarrow \omega_0 \text{ si } \max_{j=1, M} [\Pr(\omega_j | \mathbf{x})] < 1 - C_r \end{cases} \quad (2.37)$$

$\omega_0$  désigne la classe des observations rejetées en ambiguïté.

Reste à définir la valeur de  $C_r$ .

Sachant que la somme des probabilités a posteriori vaut 1, dans le cas de la classification définie dans la règle de Bayes (2.35) la quantité  $\Pr(\omega_i | \mathbf{x})$  sera comprise entre  $\frac{1}{M}$  et 1. Le rejet en ambiguïté sera donc possible pour [Fukunaga, 72] :

$$0 \leq C_r \leq 1 - \frac{1}{M} \quad (2.38)$$

### 2.4.1.3 Rejet de distance

Le rejet de distance peut être exprimé en fonction de probabilités a posteriori et de la densité de mélange. En effet, la loi du vecteur  $\mathbf{x}$ , quand sa classe d'appartenance est inconnue, est donnée par la densité de mélange  $f(\mathbf{x})$ . Ainsi tout vecteur  $\mathbf{x}$  sera affecté à la classe de rejet de distance  $\omega_d$  si sa densité de mélange est inférieure à un seuil de densité  $C_d$ .

$$\mathbf{x} \rightarrow \omega_d \text{ si } f(\mathbf{x}) = \sum_{j=1}^M \Pr(\omega_j) f(\mathbf{x} | \omega_j) < C_d \quad (2.39)$$

Plus  $C_d$  est grand, plus le rejet de distance est important. Cette valeur peut être fixée de manière heuristique à partir, par exemple, de l'ensemble d'apprentissage. En effet la densité de mélange pouvant être calculée pour tout vecteur de l'ensemble d'apprentissage, il suffit de prendre pour  $C_d$  une valeur inférieure au minimum des densités de mélange obtenues :

$$C_d \leq \min_{\mathbf{x}} f(\mathbf{x}) \quad (2.40)$$

Dans le cas paramétrique, il est assez courant d'émettre l'hypothèse selon laquelle les classes obéissent à des lois de Gauss multidimensionnelles. Les valeurs des probabilités a priori et des densités de probabilité peuvent alors être directement calculées.

On dit alors que tout vecteur  $\mathbf{x}$  de  $\mathbb{R}^d$  obéit à une loi de Gauss dans la classe  $\omega_i$  si sa

densité de probabilité s'écrit [Dubuisson et al., 01] :

$$f(\mathbf{x}|\omega_i) = (2\pi)^{-d/2} |\mathbf{S}_i|^{-1} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m}_i)^t \mathbf{S}_i^{-1} (\mathbf{x} - \mathbf{m}_i) \right\} \quad (2.41)$$

Où  $\mathbf{m}_i$  et  $\mathbf{S}_i$  sont respectivement le vecteur moyenne et la matrice de variance-covariance de la classe  $\omega_i$ .

Nous noterons que cette expression est fonction du déterminant de la matrice de variance – covariance de la classe  $\omega_i$ .

Selon [Fukunaga, 84], si les options de rejet ne sont pas appliquées, le vecteur  $\mathbf{x}$  est associé à la classe  $\omega_i$  si :

$$\begin{aligned} \ln(\Pr(\omega_i)) + \ln \left( |\mathbf{S}_i|^{-1} \right) - \frac{1}{2} (\mathbf{x} - \mathbf{m}_i)^t \mathbf{S}_i^{-1} (\mathbf{x} - \mathbf{m}_i) > \\ \ln(\Pr(\omega_j)) + \ln \left( |\mathbf{S}_j|^{-1} \right) - \frac{1}{2} (\mathbf{x} - \mathbf{m}_j)^t \mathbf{S}_j^{-1} (\mathbf{x} - \mathbf{m}_j) \quad \forall j \neq i \end{aligned} \quad (2.42)$$

$\mathbf{S}_i$  s'exprime par :

$$\mathbf{S}_i = E[(\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^t] \quad (2.43)$$

Etant donné qu'elle traduit la dispersion des observations dans la classe, elle peut aussi s'écrire autrement selon l'expression suivante :

$$\mathbf{S}_i = \frac{1}{N_i} \sum_{k=1}^{N_i} (\mathbf{x}_k - \mathbf{m}_i)(\mathbf{x}_k - \mathbf{m}_i)^t \quad (2.44)$$

Où  $N_i$  est le nombre d'échantillons de la classe  $\omega_i$ .

Le vecteur moyenne  $\mathbf{m}_i$  vaut :

$$\mathbf{m}_i = \frac{1}{N_i} \sum_{k=1}^{N_i} \mathbf{x}_k \quad (2.45)$$

Et les probabilités a priori sont simplement définies par :

$$\Pr(\omega_i) = \frac{N_i}{N} \quad (2.46)$$

Sous l'hypothèse Gaussienne, ces valeurs directes des densités de probabilité et des probabilités a priori peuvent faciliter l'application de la règle de Bayes avec options de rejet définie en (2.36), (2.37) et (2.39).

## 2.4.2. Méthodes non paramétriques

Il a été vu précédemment que les méthodes paramétriques étaient basées sur l'existence des lois de probabilités régissant les observations et les classes. Toutefois si cette connaissance est incomplète, il est préférable d'orienter la procédure de décision vers l'estimation de la loi de probabilité ou de celle des probabilités a posteriori.

Parmi les estimateurs usuels, on distingue l'estimateur de Parzen et l'estimateur des  $k$  - plus proches voisins ( $k$  - ppv).

### 2.4.2.1 Principe des méthodes non paramétriques

Considérons l'ensemble d'apprentissage constitué des  $N$  vecteurs indépendants  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ . Ces vecteurs obéissent à la même loi  $f(\mathbf{x}|\omega_i)$ .

Pour estimer cette densité de probabilité au point  $\mathbf{x}$  à partir de l'ensemble d'apprentissage, une zone  $D$  de volume  $V$  est délimitée autour du point. Ainsi en comparant le nombre  $k$  d'échantillons appartenant à  $D$  par rapport à l'ensemble des  $N$  échantillons, on démontre que l'estimation de  $f(\mathbf{x}|\omega_i)$  prend la forme :

$$\hat{f}(\mathbf{x}|\omega_i) = \frac{k}{N \cdot V} \quad (2.47)$$

Toutefois les conditions suivantes doivent être vérifiées pour que la convergence de  $\hat{f}(\mathbf{x}|\omega_i)$  vers  $f(\mathbf{x}|\omega_i)$  soit assurée :

$$\begin{aligned} \lim_{N \rightarrow \infty} V &= 0 \\ \lim_{N \rightarrow \infty} k &= \infty \\ \lim_{N \rightarrow \infty} \frac{k}{N} &= 0 \end{aligned} \quad (2.48)$$

Le choix du couple  $(k, V)$  permettra de réaliser ces trois conditions. Deux approches sont possibles :

- les différentes zones sont définies, chacune de même volume  $V$  fonction de  $N$  (par exemple  $V = \frac{1}{\sqrt{N}}$  de manière à remplir la première condition). Il suffit alors de s'assurer que la valeur de  $k$  permet de vérifier les deux autres conditions. C'est la méthode utilisée par l'estimateur de Parzen.



- le nombre d'échantillons  $k$  est choisi en fonction de  $N$  (par exemple  $k = \sqrt{N}$ ). Le volume  $V$  est ajusté, autour du point  $\mathbf{x}$ , pour que le domaine  $D$  contienne exactement ces  $k$  échantillons. C'est la méthode des  $k$  - plus proches voisins ( $k$  - ppv).

Cette dernière est présentée plus en détail dans le paragraphe suivant. C'est la première approche utilisée dans le cadre de ces travaux. Elle est très utilisée dans le cadre des méthodes non paramétriques et suscite encore beaucoup d'intérêt.

### 2.4.2.2 Règle des $k$ – plus proches voisins ( $k$ - ppv)

Reprenons la règle de classification définie en (2. 36) en l'appliquant dans un premier temps à une décision entre deux classes  $\omega_1$  et  $\omega_2$ . Ainsi :

$$\mathbf{x} \rightarrow \omega_1 \quad \text{si} \quad \Pr(\omega_1 | \mathbf{x}) > \Pr(\omega_2 | \mathbf{x}) \quad (2. 49)$$

Il vient :

$$\mathbf{x} \rightarrow \omega_1 \quad \text{si} \quad \Pr(\omega_1) f(\mathbf{x} | \omega_1) > \Pr(\omega_2) f(\mathbf{x} | \omega_2) \quad (2. 50)$$

Les densités de probabilité  $f(\mathbf{x} | \omega_i)$  ( $i=1,2$ ) sont approchées par  $\hat{f}(\mathbf{x} | \omega_i)$  à travers l'expression (2. 46), en vertu des conditions de convergence spécifiées en (2. 47). De plus les probabilités a priori des classes valent :

$$\Pr(\omega_i) = \frac{N_i}{N} \quad (2. 51)$$

Parmi les  $k$  échantillons contenus dans le volume  $V$ , il y en a  $k_1$  qui appartiennent à la classe  $\omega_1$  et les  $k_2$  restants appartiennent à la classe  $\omega_2$  (figure 2.13). La règle de décision (2. 50) devient :

$$\mathbf{x} \rightarrow \omega_1 \quad \text{si} \quad \frac{N_1}{N} \frac{k_1}{N_1 V} > \frac{N_2}{N} \frac{k_2}{N_2 V} \quad (2. 52)$$

Et par conséquent :

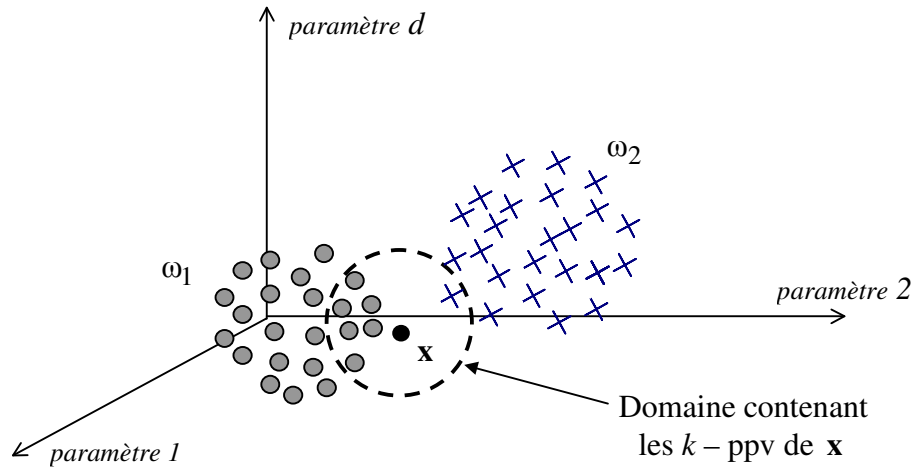
$$\mathbf{x} \rightarrow \omega_1 \quad \text{si} \quad k_1 > k_2 \quad (2. 53)$$

Cette règle peut être généralisée pour une décision entre  $M$  classes, avec  $M > 2$ .

En réalité, la manière la plus simple d'élaborer cette règle est de mesurer la distance entre le nouveau vecteur  $\mathbf{x}$  et chacun des vecteurs de l'ensemble d'apprentissage. Le vecteur  $\mathbf{x}$  sera

alors affecté à la classe la mieux représentée parmi ses  $k$  observations  $\mathbf{x}'$  les plus proches.

$$d(\mathbf{x}, \mathbf{x}')$$
(2. 54)



**Fig. 2.13.** Illustration de l'affectation par la règle des  $k$  - ppv, avec  $k = 7$

L'expression (2. 55) fournit en réalité un indice de proximité entre les deux vecteurs. Il en existe plusieurs types. Parmi les plus connus, dans le cas des données numériques, la métrique de Minkowski est très utilisée. Elle fournit un indice de "dissimilation" entre deux vecteurs  $\mathbf{x}$  et  $\mathbf{y}$  :

$$d(\mathbf{x}, \mathbf{y}) = \left[ \sum_{j=1}^d |x_j - y_j|^r \right]^{1/r} \quad r \geq 1$$
(2. 55)

Où  $d$  désigne la dimension des vecteurs.

La métrique de Minkowski possède les propriétés d'une mesure de distance car elle vérifie :

- $d(\mathbf{x}, \mathbf{y}) = 0$  si et seulement si  $\mathbf{x} = \mathbf{y}$
  - $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{y}, \mathbf{z})$
- (2. 56)

Les distances les plus utilisées pour certaines valeurs de  $r$  sont :

- La distance euclidienne classique pour  $r = 2$  :

$$d(\mathbf{x}, \mathbf{y}) = \left[ \sum_{j=1}^d |x_j - y_j|^2 \right]^{1/2} = \left[ (\mathbf{x} - \mathbf{y})^t (\mathbf{x} - \mathbf{y}) \right]^{1/2}$$
(2. 57)

- La distance de Manhattan ou city - block pour  $r = 1$  :

$$d(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^d |x_j - y_j| \quad (2. 58)$$

- Ou encore, la distance du "Max" pour  $r \rightarrow \infty$  :

$$d(\mathbf{x}, \mathbf{y}) = \text{Max}_{1 \leq j \leq d} |x_j - y_j| \quad (2. 59)$$

Enfin, il existe une mesure particulière de distance appelée distance de Mahalanobis. Elle est définie par :

$$d^2(\mathbf{x}, \mathbf{y}) = (\mathbf{x} - \mathbf{y})^t \mathbf{S}^{-1} (\mathbf{x} - \mathbf{y}) \quad (2. 60)$$

L'introduction de l'inverse de la matrice de variance – covariance  $\mathbf{S}$  dans l'expression (2. 60) signifie que cette mesure de distance prend en compte la dispersion de l'ensemble d'apprentissage, dans chaque dimension.

Bien d'autres expressions de distance peuvent être utilisées. On pourra se référer à [Wilson, 97].

L'introduction des rejets d'ambiguïté et de distance est nécessaire dans la règle des  $k$  – plus proches voisins afin de diminuer les erreurs de classifications et donc les erreurs de diagnostic.

### 2.4.2.3 Rejet d'ambiguïté

Le rejet d'ambiguïté fut introduit dans la règle des  $k$  – plus proches voisins par [Hellman, 70] et [Tomek, 76] en fixant un nombre minimal,  $k'$ , de voisins que doit avoir le vecteur  $\mathbf{x}$  dans la classe  $\omega_i$  la mieux représentée :

$$\begin{cases} \mathbf{x} \rightarrow \omega_i & \text{si } k_i = \max_{j=1, M} k_j \geq k' \\ \mathbf{x} \rightarrow \omega_0 & \text{si } k_i = \max_{j=1, M} k_j < k' \end{cases} \quad (2. 61)$$

Plus  $k'$  est proche de  $k$ , plus le rejet en ambiguïté est important. La règle de rejet d'ambiguïté a été par la suite généralisée par [Devijver, 77] en fixant un nombre minimal de voisin propre à chacune des classes.

D'autre part si  $k' = 0$ , alors le rejet n'est pas appliqué.

#### 2.4.2.4 Rejet de distance

L'option de rejet de distance peut être ajoutée à la règle de décision en fixant un seuil de distance  $T_d$ . Le seuil de distance est défini par l'utilisateur. Il est comparé par exemple :

- à la distance moyenne de  $\mathbf{x}$  à ses  $k$  plus proches voisins  $\mathbf{y}_j$  ( $j=1, k$ ) :

$$\mathbf{x} \rightarrow \omega_d \quad \text{si} \quad \frac{1}{k} \sum_{j=1}^k d(\mathbf{x}, \mathbf{y}_j) > T_d \quad (2.62)$$

- ou à la distance entre  $\mathbf{x}$  et le centre de gravité de sa classe d'affectation :

$$\mathbf{x} \rightarrow \omega_d \quad \text{si} \quad d(\mathbf{x}, \mathbf{m}_i) > T_d \quad (2.63)$$

$\mathbf{m}_i$  étant le centre de gravité de la classe d'affectation de  $\mathbf{x}$

$d(\mathbf{x}, \omega_i)$  désignera soit la distance moyenne aux plus proches voisins (2.61), soit la distance au centre de gravité (2.62).

Finalement, la règle des  $k$  – ppv incluant les deux options de rejet peut s'exprimer par :

$$\begin{cases} \mathbf{x} \rightarrow \omega_d & \text{si} \quad d(\mathbf{x}, \omega_i) > T_d \\ \mathbf{x} \rightarrow \omega_0 & \text{si} \quad d(\mathbf{x}, \omega_i) \leq T_d \quad \text{et} \quad k_i = \max_{j=1, M} k_j < k' \\ \mathbf{x} \rightarrow \omega_i & \text{si} \quad d(\mathbf{x}, \omega_i) \leq T_d \quad \text{et} \quad k_i = \max_{j=1, M} k_j \geq k' \end{cases} \quad (2.64)$$

#### 2.4.3. Calcul direct des frontières

Cette approche a pour but de fournir les expressions mathématiques des frontières séparant au mieux les classes, à l'aide des données de l'ensemble d'apprentissage. Le choix de la frontière la plus appropriée dépend de la complexité de la frontière de décision, autrement dit de la bonne ou mauvaise séparation des classes de l'ensemble d'apprentissage.

En pratique, les fonctions linéaires sont plus simples à mettre en œuvre. Mais si après une procédure de test le taux d'erreur s'avère non satisfaisant, des fonctions mathématiques plus complexes, comme par exemple les fonctions quadratiques, sont à prendre en compte.

La différence fondamentale avec une approche statistique est la possibilité, lors de la décision, d'oublier les échantillons de l'ensemble d'apprentissage au profit de l'équation des frontières.

### 2.4.3.1 Présentation du cas linéaire

Une fonction discriminante linéaire, appliquée au vecteur forme  $\mathbf{x}$ , s'écrit :

$$g_i(\mathbf{x}) = \sum_{j=1}^d w_{ij} \cdot x_j + w_{i0} = \mathbf{w}_i^t \mathbf{x} + w_{i0} \quad i = 1, \dots, M \quad (2.65)$$

Le vecteur  $\mathbf{w}_i$  est appelé vecteur – poids, et  $w_{i0}$  est une valeur de seuil. Cette dernière joue un rôle particulier dans la formulation de la fonction (2.65) surtout, et nous le verrons par la suite, du point de vue géométrique. Il est important dans cette approche d'estimer les composantes des vecteurs – poids et les seuils, pour établir la règle de décision :

$$\mathbf{x} \rightarrow \omega_i \quad \text{si} \quad g_i(\mathbf{x}) \geq g_j(\mathbf{x}) \quad \forall j = 1, M \quad (2.66)$$

Les frontières de décision linéaires sont des hyperplans séparant les régions de l'espace où sont situées les classes. Considérons deux classes  $\omega_i$  et  $\omega_j$ . La frontière séparant les zones correspondantes  $\Omega_i$  et  $\Omega_j$ , est l'hyperplan  $H_{ij}$  pour lequel :

$$g_i(\mathbf{x}) = g_j(\mathbf{x}) \quad \forall \mathbf{x} \in H_{ij} \quad (2.67)$$

En conséquence,  $H_{ij}$  a pour équation :

$$(\mathbf{w}_i^t - \mathbf{w}_j^t) \mathbf{x} + (w_{i0} - w_{j0}) = 0 \quad (2.68)$$

Que nous pouvons écrire :

$$\mathbf{w}_{ij}^t \mathbf{x} + w_{ij0} = 0 \quad (2.69)$$

Ainsi, il suffit de rechercher un seul vecteur poids et un seul seuil pour définir l'hyperplan séparant deux classes  $\omega_i$  et  $\omega_j$ .

Considérons deux classes génériques  $\omega_1$  et  $\omega_2$  constituant un ensemble d'apprentissage. En vertu de (2.68) la règle de décision est alors la suivante :

$$\begin{cases} \mathbf{x} \rightarrow \omega_1 & \text{si} \quad g(\mathbf{x}) > 0 \\ \mathbf{x} \rightarrow \omega_2 & \text{si} \quad g(\mathbf{x}) < 0 \end{cases} \quad (2.70)$$

La figure 2.14 illustre le rôle du vecteur  $\mathbf{w}$  et de  $w_0$  dans la construction de l'hyperplan séparant les deux classes.

L'hyperplan a sa normale colinéaire à  $\mathbf{w}$ , et la distance de l'origine à l'hyperplan est égale à

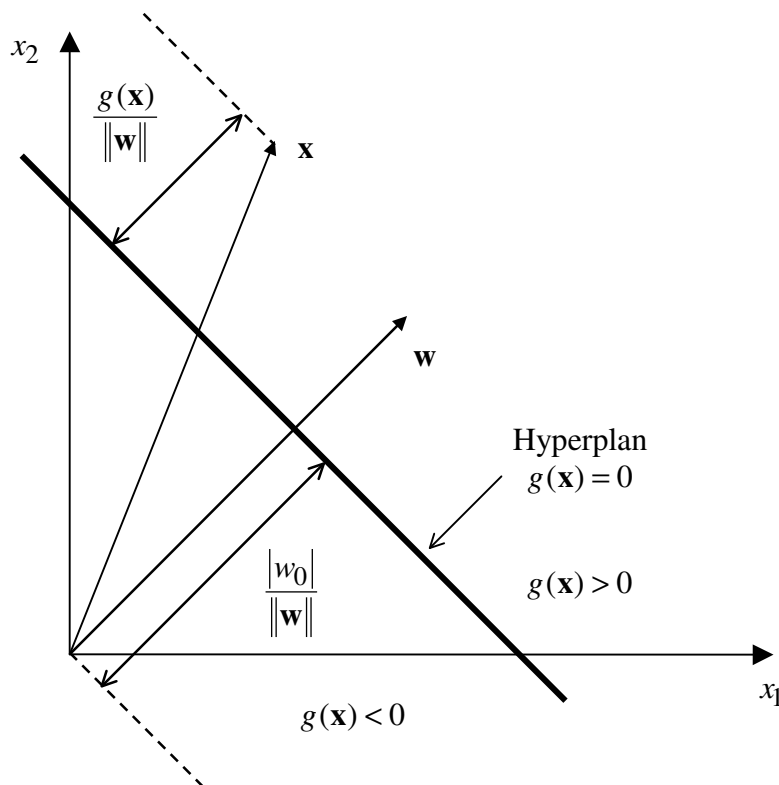
$\frac{|w_0|}{\|\mathbf{w}\|}$ . Ainsi, estimer  $\mathbf{w}$  et  $w_0$  reviendra, d'une part, à trouver l'orientation de la frontière et, d'autre part, à la situer dans l'espace. Enfin, la distance de  $\mathbf{x}$  à l'hyperplan vaut  $\frac{|g(\mathbf{x})|}{\|\mathbf{w}\|}$ .

Il est assez courant de travailler avec les expressions étendues des vecteurs formes et des vecteurs – poids :

$$\begin{aligned} \mathbf{y} &= (x_1, x_2, \dots, x_d, 1)^t = [\mathbf{x} \quad 1]^t \\ \mathbf{W} &= (w_1, w_2, \dots, w_d, w_0)^t = [\mathbf{w} \quad w_0]^t \end{aligned} \tag{2.71}$$

$g(\mathbf{x})$  est alors donné par :

$$g(\mathbf{x}) = \mathbf{W}^t \mathbf{y} \tag{2.72}$$



**Fig. 2.14.** Géométrie de la fonction discriminante linéaire dans un espace de dimension 2

Soit une quantité  $\beta$  vers laquelle on souhaite voir tendre l'expression (2.72) pour un vecteur solution, de telle sorte que :

$$\begin{cases} \beta > 0 & \text{si } \mathbf{x} \in \omega_1 \\ \beta < 0 & \text{si } \mathbf{x} \in \omega_2 \end{cases} \quad (2.73)$$

Un moyen d'avoir une bonne estimation de  $\mathbf{W}$  est d'utiliser les échantillons de l'ensemble d'apprentissage. Il faut donc définir autant de quantités  $\beta_k$  que d'échantillons  $\mathbf{x}_k$  disponibles et rechercher un vecteur  $\mathbf{W}$  tel que :

$$\mathbf{W}^t \mathbf{y}_k = \beta_k \quad k = 1, \dots, N \quad (2.74)$$

En prenant en compte les  $N$  échantillons, l'équation (2.72) peut s'écrire sous forme matricielle :

$$\mathbf{Y}\mathbf{W} = \mathbf{B} \quad (2.75)$$

Avec  $\mathbf{Y}$  la matrice de dimension  $N \times (d+1)$  et  $\mathbf{B}$  le vecteur défini par :

$$\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N)^t \quad \text{et} \quad \mathbf{B} = (\beta_1, \beta_2, \dots, \beta_N)^t \quad (2.76)$$

La résolution du système (2.74) est basée sur la minimisation d'une fonction coût  $J(\mathbf{W})$  traduisant l'écart entre la valeur désirée et la valeur obtenue.

L'expression de la fonction  $J(\mathbf{W})$  correspondant à l'erreur quadratique suivante :

$$J(\mathbf{W}) = \|\mathbf{Y}\mathbf{W} - \mathbf{B}\|^2 \quad (2.77)$$

Elle peut être minimisée en annulant son gradient :

$$\nabla J(\mathbf{W}) = 2\mathbf{Y}^t (\mathbf{Y}\mathbf{W} - \mathbf{B}) = 0 \quad (2.78)$$

On obtient ainsi, par les moindres carrés une expression estimée de  $\mathbf{W}$ , en supposant que la matrice  $(\mathbf{Y}^t \mathbf{Y})$  soit inversible :

$$\mathbf{W} = (\mathbf{Y}^t \mathbf{Y})^{-1} \mathbf{Y}^t \mathbf{B} \quad (2.79)$$

Il existe bien d'autres approches pour estimer les composantes de  $\mathbf{W}$ . En particulier, les méthodes utilisant séquentiellement les vecteurs de l'ensemble d'apprentissage sont souvent utilisées. D'autres méthodes d'élaboration de frontières de décision ont été utilisées ou énoncées dans [Didelet, 92].

### 2.4.3.2 Discriminateur linéaire de Fisher

Un cas particulier de séparateur linéaire est le discriminateur de Fisher dont l'élaboration est présentée dans [Koford, 66]. Il est basé sur la maximisation du critère de Fisher défini, pour chacun des paramètres du vecteur forme, dans le paragraphe (2.2.2.2).

Rappelons que ce critère tient compte à la fois de l'éloignement entre les classes et de leur compacité.

Le but de l'approche est de rechercher le vecteur  $\mathbf{w}$  pour lequel le critère est maximal. Prenons la formulation suivante :

$$J(\mathbf{w}) = \frac{\mathbf{w}^t \mathbf{S}_B \mathbf{w}}{\mathbf{w}^t \mathbf{S}_w \mathbf{w}} \quad (2. 80)$$

Le critère étant dérivable par rapport aux composantes de  $\mathbf{w}$ , sa valeur optimale sera obtenue en annulant son gradient. On obtient ainsi une solution directe pour  $\mathbf{w}$  :

$$\mathbf{w} = \mathbf{S}_w^{-1} (\mathbf{m}_1 - \mathbf{m}_2) \quad (2. 81)$$

Avec  $\mathbf{S}_w$  et  $\mathbf{S}_B$  les matrices de dispersion intra – classe et inter – classes (2. 14) (2. 15), calculées pour les deux classes entre lesquelles la frontière séparatrice est recherchée.

Etant donné que le centre de gravité  $\mathbf{m}$  de l'ensemble est situé sur la frontière, alors la valeur du seuil  $w_0$  peut être directement obtenue :

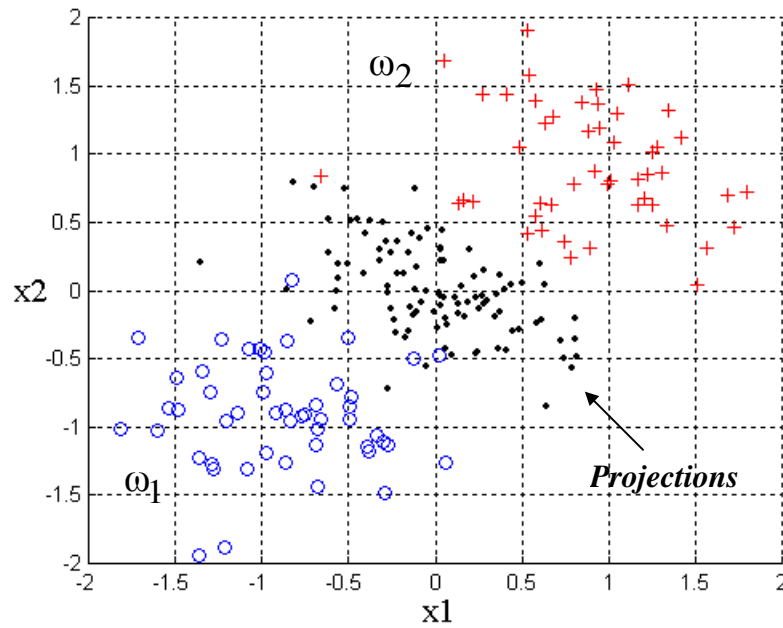
$$w_0 = -\mathbf{m}^t \mathbf{w} \quad (2. 82)$$

Avec :

$$\mathbf{m} = \frac{N_1 \mathbf{m}_1 + N_2 \mathbf{m}_2}{N} \quad (2. 83)$$

Reprenons comme exemple d'illustration l'ensemble d'apprentissage des variables de dimension 3 présenté dans le paragraphe (2.3.2) et, issu de la coalescence stricte HCM. Tous les points ont été projetés sur la frontière de décision obtenue (de normale  $\mathbf{w}$ ) afin de matérialiser celle-ci (figures 2.15 et 2.16). La séparation réalisée par la frontière est nettement plus visible dans la représentation en dimension 3.





**Fig. 2.15.** Matérialisation de la frontière de décision entre les deux classes par la projection des points – représentation en dimension 2

Le vecteur  $\mathbf{w}$  obtenu a pour composantes :

$$\mathbf{w}_{Fisher} = [-9 \ -13 \ -10.3]^t$$

Etant donné que le centre de gravité  $\mathbf{m}$  est situé en  $(0, 0, 0)$ , le seuil  $w_0$  est nul. D'autre part, les matrices de dispersions sont les suivantes :

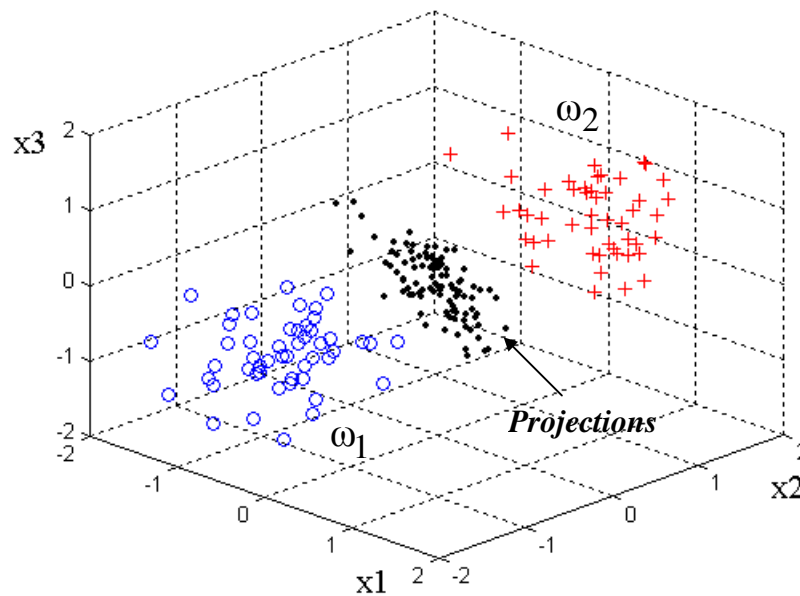
$$\mathbf{S}_w = \begin{pmatrix} 0,21 & -0,024 & 0,017 \\ -0,024 & 0,16 & -0,006 \\ 0,017 & -0,006 & 0,17 \end{pmatrix} \quad \text{et} \quad \mathbf{S}_B = \begin{pmatrix} 0,78 & 0,803 & 0,8 \\ 0,803 & 0,83 & 0,82 \\ 0,8 & 0,82 & 0,82 \end{pmatrix}$$

la valeur maximale obtenue pour  $J(2.80)$  est 14,6.

A noter que le résultat similaire pourrait être obtenu avec la solution des moindres carrées (2. 79) car le vecteur  $\mathbf{w}$  obtenu est :

$$\mathbf{w}_{MC} = [-0,29 \ -0,42 \ -0,33]^t \quad \text{avec} \quad \beta = \pm 1 \quad \text{selon l'appartenance des vecteurs.}$$

Soit un rapport d'environ 0,032 entre ses composantes et celles du vecteur estimé à l'aide du critère de Fisher. Ces deux vecteurs sont donc colinéaires.



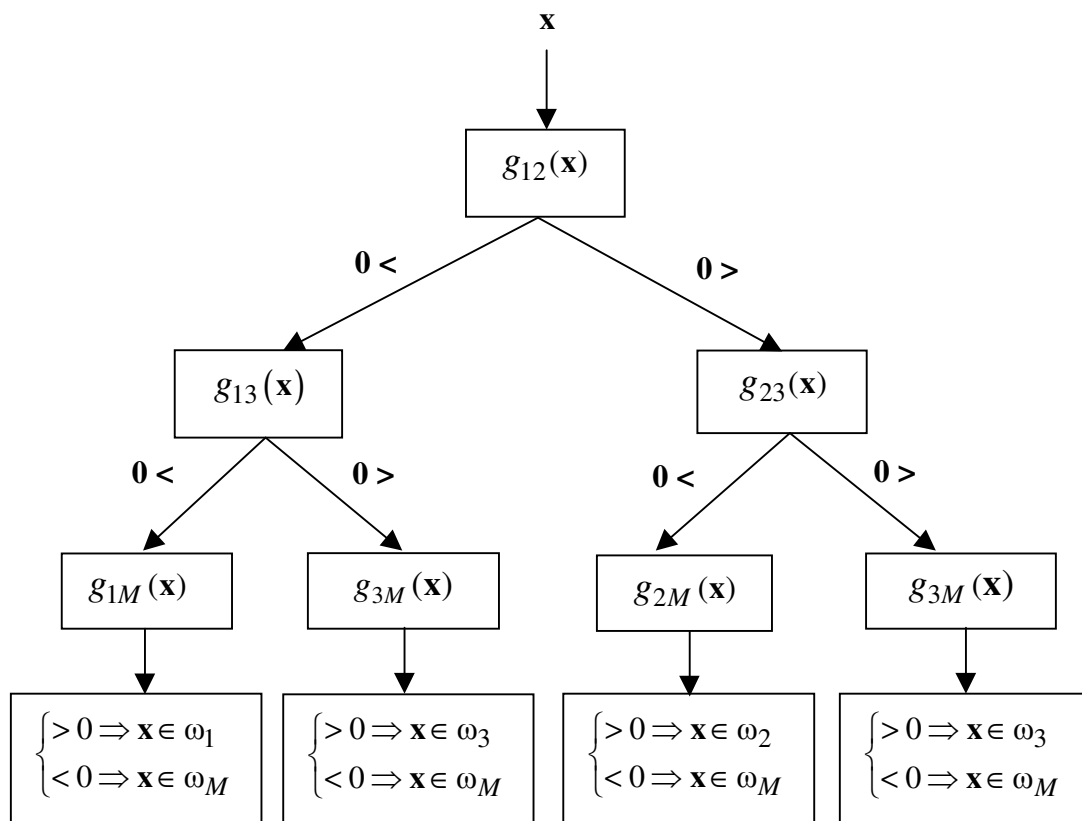
**Fig. 2.16.** Matérialisation de la frontière de décision entre les deux classes par la projection des points – représentation en dimension 3

### 2.4.3.3 Décision entre plusieurs classes

Une phase de décision en diagnostic par reconnaissance des formes est rarement limitée à une dichotomie. Au même titre que les approches statistiques, les méthodes d'analyse discriminante peuvent parfaitement être adaptées aux problèmes de décision entre plusieurs modes de fonctionnement. Le moyen le plus simple pour traiter le cas "multiclasses" est d'utiliser une approche par dichotomies successives :

$$\begin{cases} \text{si } g_{ij}(\mathbf{x}) > 0 \text{ alors tester : } g_{iu}(\mathbf{x}) \text{ avec } (u = 1, \dots, M \text{ et } u \neq j) \\ \text{si } g_{ij}(\mathbf{x}) < 0 \text{ alors tester : } g_{uj}(\mathbf{x}) \text{ avec } (u = 1, \dots, M \text{ et } u \neq i) \end{cases} \quad (2. 84)$$

Les frontières ayant été définies au préalable, il suffit d'appliquer séquentiellement la règle de décision entre deux classes  $\omega_i$  et  $\omega_j$  à toute nouvelle observation  $\mathbf{x}$ . La généralisation que nous avons utilisée est schématisée par l'arbre de décision représenté sur la figure 2.17. Chaque dichotomie testée (en accord avec (2. 84)) offre deux solutions ( $g > 0$  ou  $g < 0$ ). Ainsi,  $M - 1$  tests sont réalisés pour tout nouveau vecteur  $\mathbf{x}$ . Ce n'est que lors du test final entre la classe  $\omega_M$  et la classe ayant remporté la dichotomie précédente, que la classification de la nouvelle observation  $\mathbf{x}$  est réalisée.



**Fig. 2.17.** Généralisation de l'analyse discriminante au problème multiclassés

Les rejets d'ambiguïté et de distance peuvent aisément être introduit dans les règles de décisions utilisant les frontières :

### Rejet d'ambiguïté

On comprend intuitivement qu'aucune décision ne puisse être prise concernant un point situé à la frontière entre deux classes. Le rejet peut être appliqué en définissant une marge de sécurité  $\varepsilon$  telle que :

$$\mathbf{x} \rightarrow \omega_0 \quad \text{si} \quad |g(\mathbf{x})| \leq \varepsilon \quad (2.85)$$

### Rejet de distance

Comme pour la règle des  $k - \text{ppv}$ , il est possible de fixer un seuil pour la distance entre la nouvelle observation et sa classe d'affectation supposée.

Il a été présenté précédemment deux exemples simples de décision entre classes par calcul direct des frontières séparatrices. Toutefois il est important de noter que ces approches de discrimination linéaires ne peuvent fournir des résultats acceptables que dans le cas où les modes seraient effectivement séparables. De plus deux problèmes peuvent se poser :

- 1) il n'est pas certain qu'il existe des frontières répondant aux équations.
- 2) si une solution existe et fournit de bons résultats avec l'ensemble d'apprentissage, cela ne signifie pas pour autant qu'elle soit parfaitement adaptée aux nouvelles observations. Les coefficients des équations mathématiques devront alors être actualisés.

L'avantage principal du calcul direct des frontières est la possibilité de ne conserver, pour la décision, que leurs équations.

Enfin si cette approche simple, utilisant des séparatrices linéaires, s'avère insuffisante, alors des frontières de types quadratiques peuvent être utilisées. Leur expression généralise, en quelque sorte, la formulation des hyperplans [Didelet, 92] :

$$f(\mathbf{x}) = w_0 + \sum_{i=1}^d w_i x_i + \sum_{i=1}^d \sum_{j=1}^d w_{ij} x_i x_j \quad (2.86)$$

## 2.5. Evaluation des performances de la procédure

Quelles que soient les approches utilisées pour la mise en place du processus de diagnostic par reconnaissance des formes, il est primordial de justifier la robustesse du classifieur choisi. En effet des erreurs de diagnostic pourraient découler de mauvaises performances de la procédure de décision. De plus, les performances de la classification sont dépendantes (surtout en diagnostic) de la pertinence du vecteur – forme, autrement dit de la signature du système analysé.

### 2.5.1. Critères de qualité

Deux critères permettent d'évaluer les performances de la règle de décision choisie par rapport au problème posé :

- Le taux d'observations bien classées :

$$T_B = \frac{\text{Nombre de bien classées}}{N} = \frac{N_B}{N} \times 100 \quad (2.87)$$

- Et le taux de mal classées :

$$T_M = \frac{\text{Nombre de mal classées}}{N} = \frac{N_M}{N} \times 100 \quad (2.88)$$

Ces deux quantités ont beaucoup d'importance. Si taux d'erreur est proche de 0 %, alors on peut estimer que non seulement, "en amont", le vecteur - forme caractérise bien l'ensemble d'apprentissage, mais aussi que la règle de décision est adaptée. Parallèlement, il est souhaitable d'avoir une valeur très faible de la probabilité de mauvaise classification lors de la phase de test afin d'avoir un minimum d'erreur de diagnostic concernant les nouvelles observations.

### 2.5.2. Méthodes de test

C'est à l'utilisateur de fixer les procédures de test du classifieur afin d'en vérifier les performances.

De l'ensemble des données disponibles  $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$  dont sont connues les classes d'appartenance, peut être extrait une partie  $\mathbf{X}_a$  réservée à l'apprentissage et une autre partie  $\mathbf{X}_t$  constituée de vecteurs réservés au test. Ces deux ensembles sont disjoints.

Citons deux possibilités de choix pour  $\mathbf{X}_a$  et  $\mathbf{X}_t$  :

- $\mathbf{X}_t$  est constitué d'un seul vecteur à la fois,  $\mathbf{X}_a$  contient les  $N-1$  autres vecteurs :  $\mathbf{X}_a = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{k-1}, \mathbf{x}_{k+1}, \dots, \mathbf{x}_N\}$  et  $\mathbf{X}_t = \{\mathbf{x}_k\}$ . C'est la méthode du leave one out. Cette procédure est répétée pour chacune des  $N$  partitions. A l'issue du test sur l'ensemble des éléments, les résultats sont moyennés. L'inconvénient de cette méthode est son coût calculatoire même si elle fournit une bonne estimation du taux d'erreur.
- $\mathbf{X}_t$  est constitué d'un nombre non négligeable de vecteurs, n'ayant pas servi à l'apprentissage. Cette méthode fournit une estimation plutôt pessimiste du taux d'erreur, mais a le mérite de pouvoir être lancée à plusieurs reprises en sélectionnant aléatoirement les données. Si après avoir effectué une moyenne de l'ensemble des résultats le taux de bonne classification s'avère acceptable, alors le vecteur forme et la règle de décision sont considérés comme étant robustes.

La décision de redéfinir ou non la signature et l'ensemble d'apprentissage dépend des résultats du test.

## 2.6. Exploitation du système de diagnostic

Les différentes méthodes précédemment citées (caractérisation du vecteur – forme, apprentissage, élaboration d'une règle de décision avec options de rejet) constituent les deux premières phases de la mise en place du système de diagnostic par RDF.

Il s'agira par la suite de *reconnaître* à quel mode de fonctionnement correspond une nouvelle observation effectuée à un instant donné sur le système. Il est bien entendu que, sur la base de la connaissance disponible, l'algorithme de décision ne pourra que confirmer un mode existant ou rejeter une observation jugée trop ambiguë ou différente des autres. En cas de rejet, seul un travail d'expertise peut fournir une analyse complète des causes de ce changement d'état.

La méthode doit être adaptative. Elle doit permettre de rechercher une nouvelle structure en classe pour les points rejetés, dès lors que leur nombre devient significatif. Un nouvel apprentissage est à relancer afin de prendre en compte cette classe supplémentaire. La règle de décision est aussi à adapter.

## 2.7. Conclusion

Dans ce chapitre nous avons décrit les différentes étapes de l'élaboration d'un système de diagnostic par reconnaissance des formes. Le principe général, comme pour tout autre outil de diagnostic, est de confronter les données recueillies sur le processus aux informations sur les modes de fonctionnement, disponibles dans la base de connaissance.

L'accent a été porté sur les bases théoriques des méthodes de construction du vecteur forme (à partir des paramètres disponibles) et des règles de décision.

Les différentes possibilités de règles de décision présentées se divisent en trois groupes :

- les méthodes paramétriques, utilisables lorsque les lois de probabilité, régissant les observations et les classes, sont connues. La règle de Bayes entre dans le cadre de ces méthodes.
- les méthodes non paramétriques : elles constituent une estimation des lois de probabilité ( $k$  – ppv par exemple).

- 
- Le calcul direct des frontières entre classes par le biais d'analyses discriminantes linéaires ou non linéaires.

Parmi ces différentes possibilités de choix du système de décision, la règle des  $k$  – plus proches voisins et l'analyse discriminante linéaire utilisant le critère de Fisher ont particulièrement attiré notre attention. Ces deux règles sont simples à élaborer et ne nécessitent pas au départ la connaissance des lois de probabilité.

Toutefois, nous avons vu que le diagnostic basé sur la RDF peut se heurter à un certain nombre de difficultés liées aux erreurs de classification si les modules de rejets de distance et d'ambiguïté ne sont pas ajoutés. Le but du rejet est aussi de diminuer les erreurs de classification.





---

## CHAPITRE III – Application au diagnostic de la machine asynchrone

---

### **3.1 Introduction**

Les deux chapitres précédents ont permis de présenter les méthodes classiques d'analyse de la machine asynchrone et, sur l'utilisation des méthodes de reconnaissance des formes en vue du diagnostic.

Ce chapitre porte sur l'application proprement dite de cette procédure à la machine asynchrone.

Le banc moteur et les modes de fonctionnement sont présentés au début de ce chapitre. Nous exposons dans un deuxième temps une étude complète sur la recherche des paramètres du vecteur forme. Ces paramètres sont calculés en utilisant les méthodes classiques de traitement des signaux. Les études de sensibilité de ces paramètres en fonction des différentes configurations et niveaux de charge permettent d'avoir une première idée sur leur pertinence. Elles sont réalisées sur un échantillon représentatif des divers modes de fonctionnement étudiés sur la machine.

Mais l'analyse de la variation des paramètres n'est pas suffisante pour définir le vecteur forme. Deux algorithmes de sélection de caractères sont utilisés :

- la sélection basée sur le calcul du rapport de Fisher. Cette méthode vise à ranger les paramètres dans l'ordre décroissant de leur valeur de critère (de compacité / séparabilité) correspondante.
- l'algorithme "Sequential Backward Selection" (SBS) qui est une des méthodes consistant à rechercher un sous – ensemble de  $d'$  paramètres pertinents.

Nous utilisons ainsi l'ensemble d'apprentissage au travers de deux vecteurs formes différents. Cela nous permet de proposer deux choix de signatures de la machine.

Après caractérisation de l'ensemble des classes (ou modes de fonctionnement), les deux

---

règles de décisions, à savoir la règle des  $k - ppv$  et les frontières linéaires seront appliquées aux nouvelles observations (en défaut ou non) du système. Les résultats de diagnostic obtenus sont comparés.

Il s'agit dans cette phase d'application de démontrer l'efficacité des méthodes de RDF utilisées pour suivre l'état de la machine asynchrone.

Dans un premier temps, il a été nécessaire de valider l'approche du diagnostic par RDF, avant de l'implanter sur notre système. Autrement dit, il a fallu faire une étude de faisabilité avant de passer à la réalisation. Dans ce cadre, des acquisitions ont été menées sur un nombre limité de modes de fonctionnement et de niveaux de charge, avec ou sans onduleur. Les résultats correspondants sont présentés au paragraphe 3.3 (Recherche de signatures).

Pour l'étape de réalisation du système de décision, présenté au paragraphe 3.4 (Procédure de décision), il était prévu de réaliser des acquisitions supplémentaires afin d'enrichir la base de connaissance initiale. Malheureusement, à la suite d'une avarie sur l'onduleur, et pour répondre à la demande du GDR ME2MS<sup>(1)</sup>, il a été décidé de redéfinir des modes de fonctionnement et des niveaux de charge différents. Cela ne remettait pas en cause l'étude de faisabilité, mais induisait une rupture dans notre approche progressive. Toutefois, pour bien montrer le caractère générique de l'approche RDF, nous avons préféré travailler avec deux campagnes de mesures successives.

---

<sup>1</sup> GDR ME2MS (Maîtrise de l'Energie Electrique du Matériau aux Systèmes)

## 3.2 Description du procédé

### 3.2.1 Présentation du banc moteur

Le moteur asynchrone est caractérisé par :

- La tension nominale entre phase : 400 V
- La fréquence d'alimentation : 50 Hz
- La vitesse nominale : 1440 tr/min
- La puissance utile nominale : 5.5 kW
- Le facteur de puissance :  $\cos \varphi = 0.84$
- Le courant nominal : 11,4 A
- Le nombre de paires de pôles :  $p = 2$
- La résistance stator par phase : 1,315  $\Omega$
- Le nombre d'encoches au rotor  $N_r = 28$
- Le nombre d'encoches au stator  $N_s = 48$

Les enroulements au stator sont couplés en étoile. La photographie suivante présente le banc moteur utilisé.



**Fig. 3. 1.** Photo du banc moteur 5,5 kW dédié au diagnostic.

La charge est composée d'une machine à courant continu et d'un banc de résistances. Au niveau de l'induit, la machine à courant continu peut délivrer jusqu'à 60 A sous environ 115 V, soit près de 7 kW à 2800 tr/mn. La résistance d'induit est de 214 m $\Omega$ .

### 3.2.2 Moyens de mesures

Les signaux ont été relevés, lors des différents modes de fonctionnement, à l'aide d'un système d'acquisition *Odyssey®* de marque *Gould Nicolet*.

Celui-ci possède huit entrées différentielles dites voies rapides qui peuvent être échantillonnées jusqu'à 10 MHz sur 14 bits et, huit entrées en mode commun dites voies lentes sur lesquelles la fréquence d'échantillonnage peut atteindre 1 MHz sur 16 bits. Quatre entrées sont réservées aux accéléromètres mesurant les vibrations. Les entrées "accéléromètres" offrent aussi la possibilité d'obtenir une fréquence d'échantillonnage de 100 kHz. Les accéléromètres piézo-électriques sont alimentés par des sources de courants constants fournissant entre 1 et 10 mA sous 28 V. Enfin quatre autres entrées sont destinées aux mesures de températures.

La figure 3.2 représente le schéma du dispositif : ensemble machine charge et acquisition des signaux.

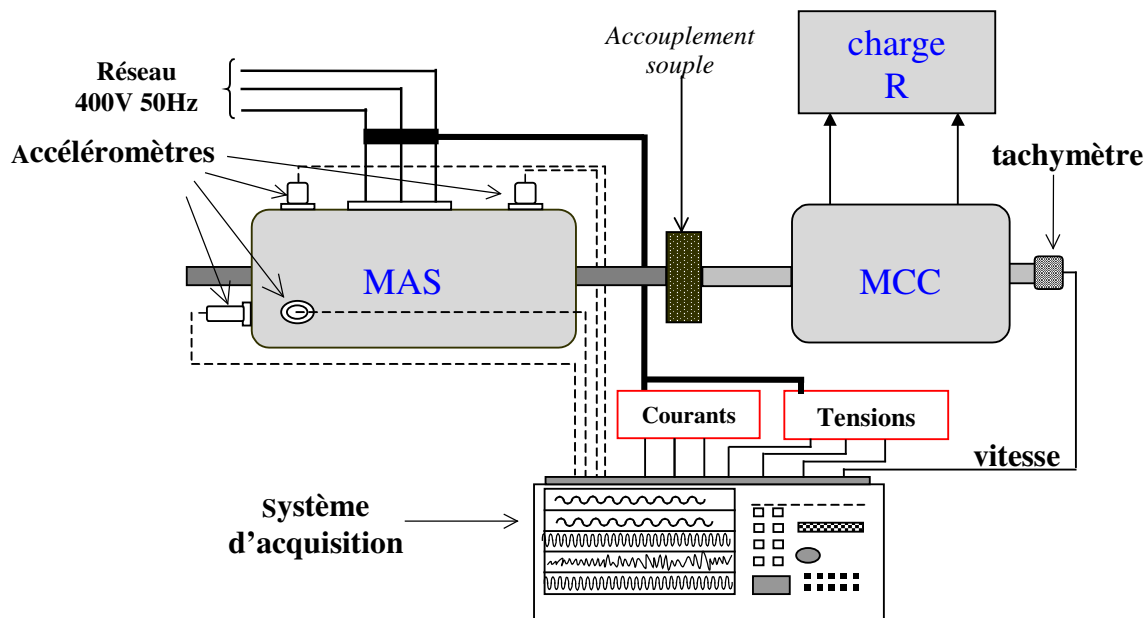


Fig. 3.2. Schéma de principe de l'acquisition des signaux

---

Les signaux enregistrés par le système d'acquisition sont :

- les trois courants d'alimentation de la machine
- les trois tensions d'alimentation de la machine
- la vitesse de rotation
- les vibrations dans les directions suivantes :
  - direction verticale avec deux accéléromètres, l'un côté opposé accouplement et l'autre côté accouplement
  - direction parallèle à l'axe du moteur, côté opposé accouplement
  - direction horizontale et perpendiculaire à l'axe du moteur, côté opposé accouplement

Toutes les acquisitions ont été réalisées en régime permanent sur une durée de 10 secondes, avec une fréquence d'échantillonnage de 10 kHz. Soit 100.000 points pour chacun des signaux mesurés. Pour le traitement (notamment fréquentiel) des signaux, il faut prendre un nombre de points en  $2^N$ . Intuitivement il faudrait choisir  $N_p = 2^{16}$  (car il s'agit du nombre le plus proche du nombre total d'échantillons sur chaque voie), mais les signaux sont filtrés, avant le calcul spectral, par un filtre passe – bas numérique tenant compte des conditions initiales. Au final, les signaux sont analysés sur  $N_p = 2^{15} = 32768$  points. Ainsi pour le calcul fréquentiel, la résolution spectrale  $\Delta f$  est de :

$$\Delta f = \frac{F_e}{N_p} = 0,3 \text{ Hz} \quad (3.1)$$

Pour chaque mode de fonctionnement, 15 enregistrements ont été réalisés.

### 3.2.3 Différents modes de fonctionnement

Les modes de fonctionnement ayant servis à valider la procédure de diagnostic sont repris dans le tableau (3.1). Chacun d'entre eux sera représenté par une classe dans l'espace de décision. Les défauts ont été créés à la fois au stator et au rotor et, pour plusieurs niveaux de charge. Ce tableau scindé en deux parties fait apparaître les deux campagnes de mesures autour desquelles s'articule ce chapitre.

<b>Mode de fonctionnement</b>	<b>Taux de charge</b> (Pour une puissance nominale : $P_u = 5,5\text{kW}$ )	<b>Alimentation</b>
Machine saine	0%, et à $P_u \approx 3\text{kW}$	Réseau et onduleur
4 barres cassées au rotor		
Déséquilibre au stator		
Machine saine	0%, 25%, 50%, 75%, 100%	Réseau
3 barres cassées au rotor		
Déséquilibre au stator	0%, 25%, 50%, 75%	
1 barre partiellement cassée au rotor		

**Tableau. 3. 1.** Les modes de fonctionnement qui seront analysés.

Les mesures "machine saine", "4 barres cassées" et "déséquilibre stator" (à vide et à 3 kW) en alimentation réseau, collectées lors de la première campagne d'acquisition ont permis d'appréhender les algorithmes de reconnaissance des formes [Casimir, 02]. Elles seront utilisées tout au long de la partie suivante pour illustrer la sensibilité des paramètres calculés et la définition du vecteur forme.

Les autres mesures correspondant aux modes de fonctionnement, en alimentation réseau et à plusieurs niveaux de charge, ont été réalisées par la suite, dans le but d'appliquer le système de décision sur un ensemble d'observations plus riche.

Au niveau des rotors, les barres ont été cassées en perçant les conducteurs des cages d'écureuil. Les rotors endommagés sont visibles sur la figure 3.3.



**Fig. 3. 3.** Photos des rotors utilisés

Le déséquilibre au stator a pu être provoqué en plaçant une résistance de  $100\text{m}\Omega$  sur une des phases. Il simule un défaut de court-circuit entre spires.

Les données obtenues sur la machine alimentée par onduleur vont permettre de montrer l'efficacité des algorithmes de décision, même dans ce type d'alimentation.

### 3.3 Recherche de signatures

La recherche de signatures consiste, en reconnaissance de formes, à construire le vecteur forme pour définir de manière pertinente les observations effectuées sur le système. En outre, le vecteur forme (qui est, à proprement parler, la signature du processus) doit permettre de bien distinguer les classes.

Cette étape de la phase d'apprentissage est primordiale. En effet, contrairement aux méthodes classiques de diagnostic, l'application de la RDF nécessite la définition d'une seule et même signature dont les paramètres varient en fonction de l'état du système.

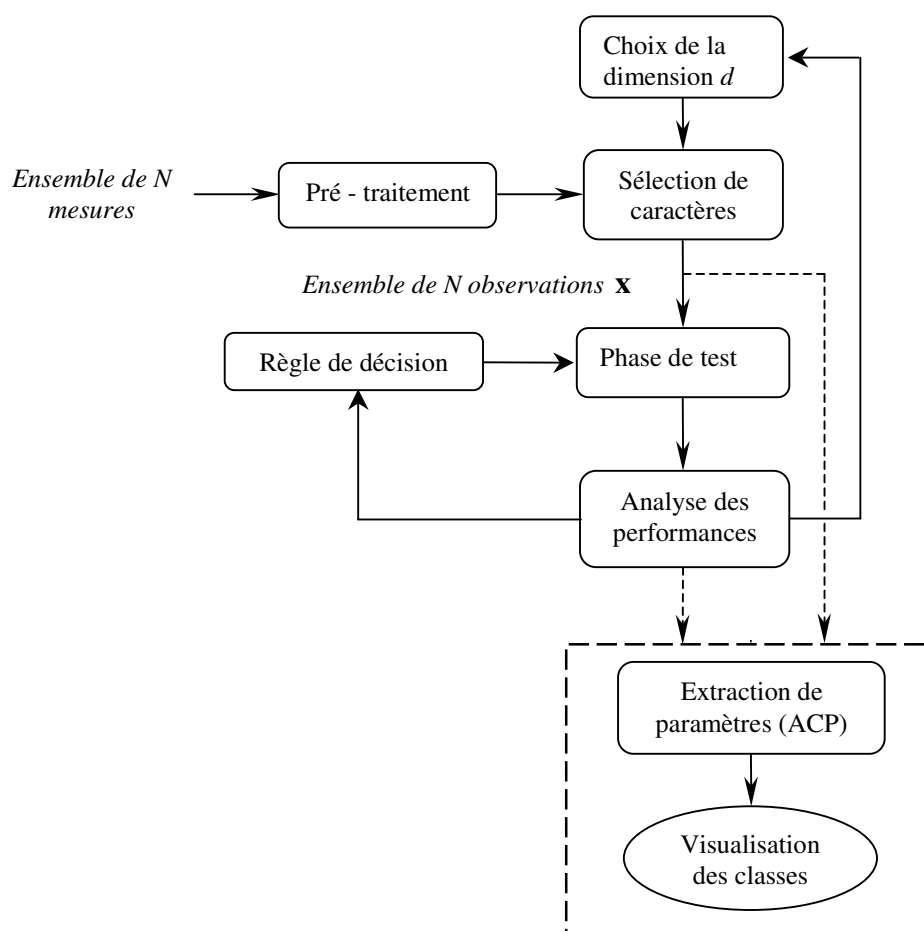
Parmi tous les paramètres calculés, et que nous allons présenter, certains risquent de ne pas être pertinents pour les modes de fonctionnement étudiés. Il faut donc soit les analyser un par un en visualisant leur courbe de sensibilité, soit utiliser des algorithmes de sélection de caractères afin de réduire la dimension de l'espace en sélectionnant les plus pertinents. La recherche de la signature par le biais d'algorithmes de sélection est beaucoup plus rapide.

Une autre méthode de réduction de la dimension est l'extraction de paramètres par analyse en composantes principales (ACP). Cette méthode offre, par ailleurs, la possibilité de visualiser la position des classes. L'ACP sera utilisée dans cet objectif. Elle permettra

simplement de voir si les classes sont bien séparées.

Toutefois, le choix du vecteur forme ne pourra être validé qu'après la phase de test du système. En effet, après l'apprentissage il est nécessaire de tester la règle de décision choisie sur une partie des données n'ayant pas servi à l'apprentissage. Ainsi, les taux de bonnes ou de mauvaises classifications seront les éléments qui permettront de voir si les paramètres du vecteur forme sont performants dans leur capacité à caractériser les différentes classes (à condition de ne pas utiliser une mauvaise règle de décision).

En résumé, la détermination de la signature sera entreprise selon le schéma de principe de la figure 3.4.



**Fig. 3. 4.** Principe de sélection de la signature (vecteur forme) utilisé dans l'application.

Cette procédure nécessite le fait que la règle de décision soit fixée. Dans le cadre de l'étude, la règle de décision choisie est la règle des  $k - ppv$  avec  $k = 5$ . Pour chaque mode de fonctionnement, 15 acquisitions sont réalisées. Nous en gardons 10 dans chaque cas pour définir l'ensemble d'apprentissage  $X_a$ . C'est sur la base de ces  $M \times 10$  vecteurs (où  $M$  désigne



---

le nombre de modes de fonctionnement ou classes) que la sélection des paramètres est effectuée. Les autres vecteurs, n'ayant pas servi à l'apprentissage, constituent l'ensemble de test  $X_t$ . Selon les performances du classifieur, la dimension de l'espace de représentation (taille du vecteur forme) est modifiée. Cette procédure n'exclut pas la possibilité de modifier également la règle de décision (le nombre de plus proches voisins par exemple).

### 3.3.1 Prétraitement des signaux et calcul des paramètres.

L'objectif du système de diagnostic est, rappelons le, de pouvoir identifier un mode déjà connu parmi les modes disponibles dans la base de données. C'est aussi, et surtout, la possibilité de détecter les nouveaux modes de fonctionnement lorsque ceux – ci apparaissent et de pouvoir les intégrer dans sa base de données après expertise.

Cela est difficilement réalisable si la recherche de signature est basée sur un seul type d'analyse. Prenons un exemple :

Il est bien connu qu'en fonctionnement en charge et en régime permanent le défaut "barre cassée" fait apparaître (ou amplifier) certaines composantes dans le spectre des courants au stator. L'approche par analyse spectrale est donc la mieux adaptée. Le problème est que non seulement les raies caractéristiques sont difficilement identifiables avec une alimentation onduleur, mais en plus elles n'existent quasiment pas à très faible charge. Enfin, la seule analyse de ces raies sur les courants pourrait être inefficace pour détecter l'apparition d'autres défauts tels que ceux qui sont liés au stator.

C'est pour cela que les paramètres du vecteur forme sont issus de divers types d'analyses. Nous allons lister, au cours des paragraphes suivants, les différents paramètres qui découlent du traitement des signaux courants et tensions et de la connaissance du fonctionnement de la machine asynchrone.

#### 3.3.1.1 Données temporelles

Les analyses ayant permis d'extraire des signaux les paramètres "candidats" au vecteur forme, vont être présentées au cours de ce paragraphe. Il s'agira par la suite d'en sélectionner les plus pertinents en fonction de l'ensemble d'apprentissage.

### 3.3.1.1.1 Transformation triphasée – diphasée (axes $\alpha$ , $\beta$ liés au stator)

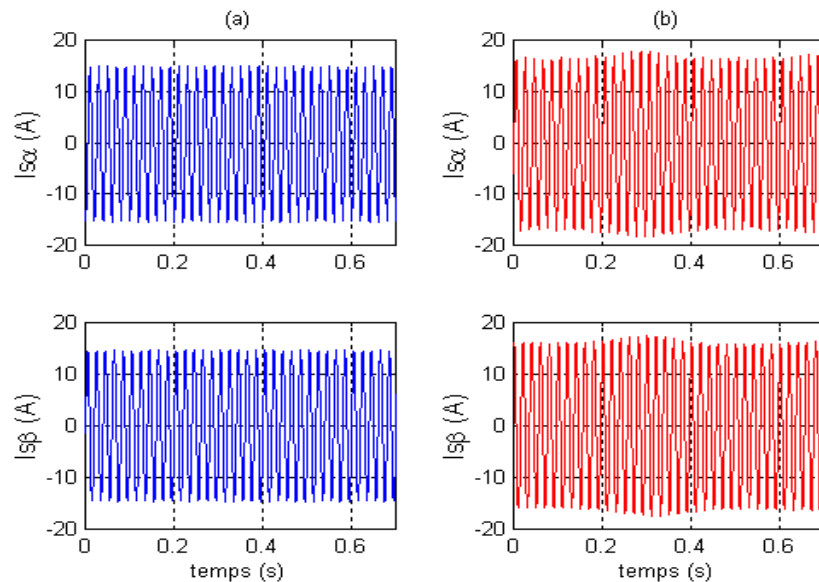
Cette transformation consiste en la substitution des enroulements triphasés (a, b, c) disposés au stator par des enroulements diphasés fictifs ( $\alpha$ ,  $\beta$ ). Avec l'axe  $O\alpha$  confondu avec l'axe  $Oa$  qui représente l'axe de symétrie de la phase a (figure 1.5).

La transformation est la même pour les flux, les tensions et les courants. Pour ces deux derniers, elle s'exprime selon l'équation suivante :

$$\begin{bmatrix} V_{s\alpha} \\ V_{s\beta} \\ V_{s0} \end{bmatrix} = [C] \cdot \begin{bmatrix} V_{sa} \\ V_{sb} \\ V_{sc} \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} I_{s\alpha} \\ I_{s\beta} \\ I_{s0} \end{bmatrix} = [C] \cdot \begin{bmatrix} i_{sa} \\ i_{sb} \\ i_{sc} \end{bmatrix} \quad (3.2)$$

[C] est la matrice de concordia définie par l'équation (1.4).

L'étude porte sur l'évolution temporelle des composantes en quadrature  $I_{s\alpha}$  et  $I_{s\beta}$ . Celles-ci sont modulées en amplitude en présence de barres cassées tout comme les courants au stator ( $i_{sa}$ ,  $i_{sb}$  et  $i_{sc}$ ). La figure 3.5 illustre l'affectation temporelle des courants en alimentation réseau, dans le cas du défaut rotor (4 barres cassées dans cet exemple) par rapport au fonctionnement sain.



**Fig. 3. 5.** Evolution des composantes  $I_{s\alpha}$  et  $I_{s\beta}$  des courants stators pour un rotor sain (a) et pour un rotor ayant 4 barres cassées (b)

### 3.3.1.1.2 Puissances

A partir des composantes en quadrature des tensions et des courants, il est aisé d'estimer les puissances actives, réactives et apparentes consommées par la machine.

Ecrivons les tensions et courants, de manière générique, sous forme complexe en considérant les variables  $\underline{v}$  et  $\underline{i}$  telles que :

$$\begin{cases} \underline{v} = V_s \alpha + j V_s \beta \\ \underline{i} = I_s \alpha + j I_s \beta \end{cases} \quad (3.3)$$

La puissance active  $P$  et la puissance réactive  $Q$  s'expriment en fonction de ces variables [Vas, 92] :

$$\begin{cases} P = \text{Re} \left[ \underline{v} \cdot \underline{i}^* \right] = V_s \alpha I_s \alpha + V_s \beta I_s \beta \\ Q = \text{Im} \left[ \underline{v} \cdot \underline{i}^* \right] = V_s \beta I_s \alpha - V_s \alpha I_s \beta \end{cases} \quad (3.4)$$

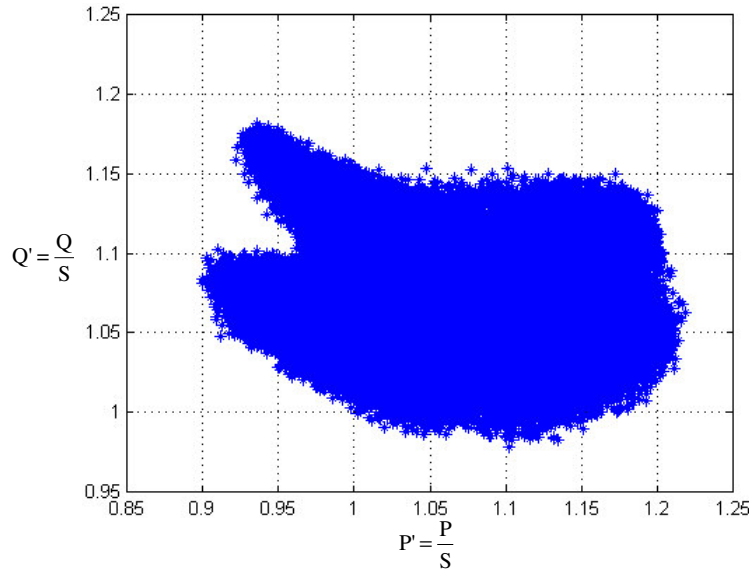
Où  $\underline{i}^*$  est l'expression conjuguée du courant. Enfin la puissance apparente complexe est :

$$\underline{S} = P + jQ \quad (3.5)$$

Mais pour la suite, nous tiendrons simplement compte de son expression en fonction de la valeur efficace des tensions et des courants :

$$S = 3 \cdot V_{\text{eff}} \cdot I_{\text{eff}} \quad (3.6)$$

Il est apparu intéressant d'analyser la caractéristique normalisée  $Q'(P')$  qui est un nuage de  $N_p$  points (figure 3.6).  $Q'$  et  $P'$  sont les puissances actives et réactives normalisées par rapport à la puissance apparente.



**Fig. 3. 6.** Caractéristique normalisée  $Q'(P')$  obtenue pour la machine saine, pour une charge de 3 kW en alimentation réseau

Les points de la caractéristique (figure 3.6) représentent de manière normalisée l'évolution de la puissance apparente consommée par la machine à chaque instant.

Trois paramètres sont calculés à l'aide de cette caractéristique.

Le premier est la dispersion du nuage de points, notée  $\zeta$ . Les deux autres sont les valeurs moyennes de  $P'$  et de  $Q'$  que nous noterons  $m_p$  et  $m_q$  :

Soit  $\mathbf{m}_{p,q}$  le centre de gravité de l'ensemble et  $\mathbf{x}_k$  un point de coordonnées  $(P'(k), Q'(k))$ .

$$\mathbf{m}_{p,q} = [m_p \ m_q] \tag{3.7}$$

Avec :

$$\begin{cases} m_p = \frac{1}{N_p} \sum_{k=1}^{N_p} P'(k) \\ m_q = \frac{1}{N_p} \sum_{k=1}^{N_p} Q'(k) \end{cases} \tag{3.8}$$

De ce fait :

$$\zeta = \sum_{k=1}^{N_p} (\mathbf{x}_k - \mathbf{m}_{p,q})(\mathbf{x}_k - \mathbf{m}_{p,q})^t \tag{3.9}$$

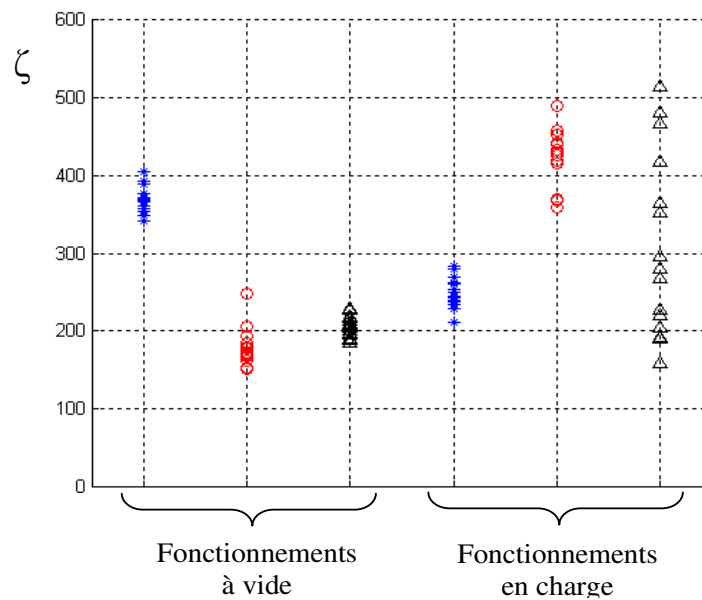
Le paramètre  $\zeta$  est un indicateur de dispersion car il s'agit de la somme des distances au

carré entre chaque point de la caractéristique et le centre de gravité.

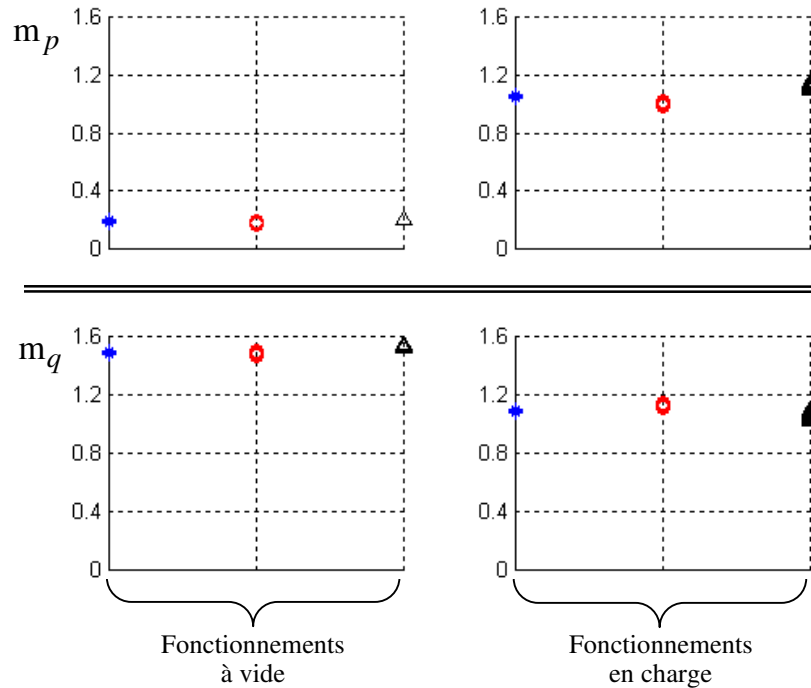
Afin de juger de l'intérêt des paramètres  $m_p$ ,  $m_q$  et  $\zeta$ , prenons pour base 15 acquisitions de chacun des modes de fonctionnements suivants :

- Machine saine à vide
- Machine saine en charge
- 4 barres cassées à vide
- 4 barres cassées en charge
- Déséquilibre au stator à vide
- Déséquilibre au stator en charge

Les figures 3.7 et 3.8 permettent de suivre la variation de ces paramètres sur l'ensemble des 15 acquisitions.



**Fig. 3. 7.** Variations de la dispersion  $\zeta$  sur l'ensemble des acquisitions à vide et en charge. Machine saine (\*). Rotor en défaut (o). Déséquilibre stator ( $\Delta$ )



**Fig. 3. 8.** Variations des puissances moyennes. Machine saine (\*). Rotor en défaut (o). Déséquilibre stator ( $\Delta$ )

La première observation qui peut être faite, est que le paramètre de dispersion est relativement sensible aux différents états de la machine. Cela est sans doute dû à l'affectation temporelle des courants.

D'un autre côté, l'évolution des paramètres de puissance moyenne est corrélée avec le niveau de charge. Ces paramètres, à eux seuls, risquent de ne pas permettre une distinction franche entre les différents modes, en tout cas du point de vue des défaillances.

Mais combinés à d'autres paramètres sensibles, ces deux paramètres pourraient s'avérer pertinents. En effet ils rendent compte de la puissance mise en jeu, ce qui complètera l'information fournie par les autres paramètres.

### 3.3.1.1.3 Calcul sur $I_{s\alpha}$ et $I_{s\beta}$

Outre le calcul de puissance, d'autres analyses ont été réalisées à partir des courants en quadrature pour extraire plus de paramètres. Les paramètres suivants sont directement liés à l'évolution temporelle des signaux et à la représentation dans le plan de phase du courant stator. Ils sont cinq au total.

Considérons le module du courant stator, noté simplement  $I_{s\alpha,\beta}$ , en fonction des

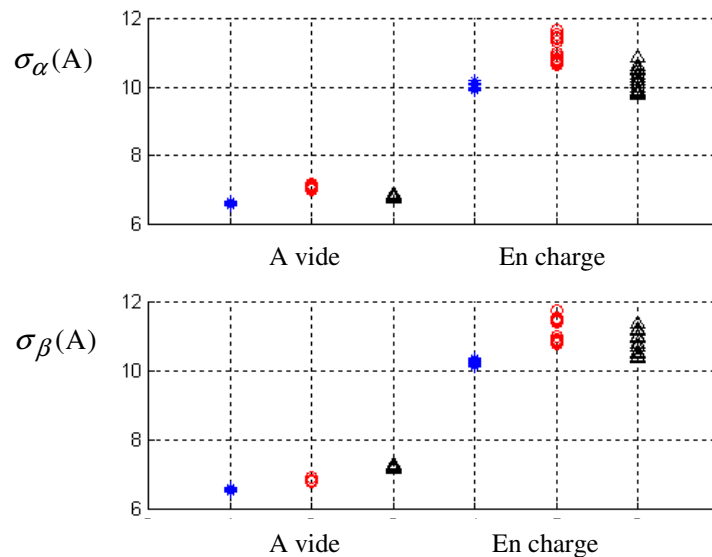
composantes en quadratures (3.10).

$$I_{s\alpha,\beta} = \sqrt{I_{s\alpha}^2 + I_{s\beta}^2} \quad (3.10)$$

On dispose ainsi de trois signaux dont il s'agit dans un premier temps de calculer les écarts–types, à l'aide de leurs valeurs moyennes notées respectivement  $m_\alpha$ ,  $m_\beta$  et  $m_{\alpha,\beta}$  :

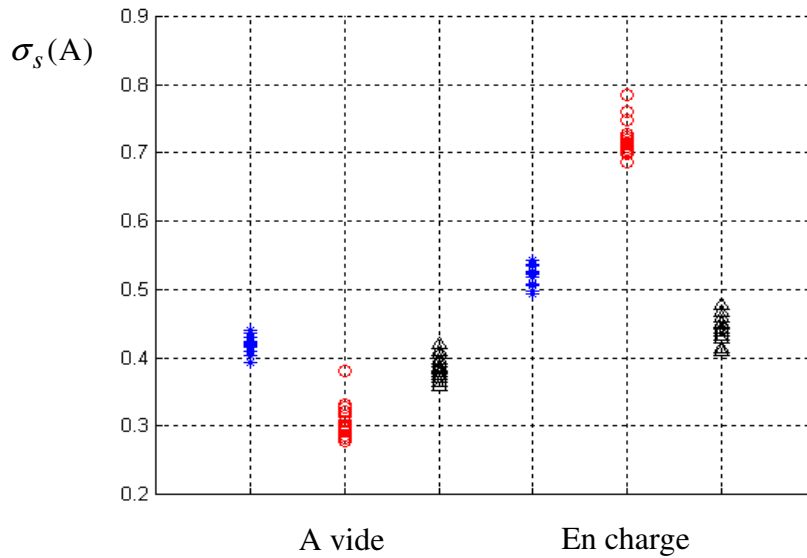
$$\begin{cases} \sigma_\alpha = \left[ \frac{1}{N_p} \sum_{k=1}^{N_p} (I_{s\alpha}(k) - m_\alpha)^2 \right]^{1/2} \\ \sigma_\beta = \left[ \frac{1}{N_p} \sum_{k=1}^{N_p} (I_{s\beta}(k) - m_\beta)^2 \right]^{1/2} \\ \sigma_s = \left[ \frac{1}{N_p} \sum_{k=1}^{N_p} (I_{s\alpha,\beta}(k) - m_{\alpha,\beta})^2 \right]^{1/2} \end{cases} \quad (3.11)$$

Sur l'ensemble des acquisitions, à vide et en charge, les paramètres  $\sigma_\alpha$  et  $\sigma_\beta$  (figure 3.9) fournissent un résultat quasi identique, à savoir une bonne distinction du fonctionnement à vide par rapport au fonctionnement en charge, mais un faible niveau de détection des défauts par rapport au fonctionnement sain.



**Fig. 3. 9.** Variations des écarts – types des courants  $I_{s\alpha}$  et  $I_{s\beta}$ . Machine saine (\*). Rotor en défaut (o). Déséquilibre stator ( $\Delta$ )

Sur la figure 3.10, nous constatons que le paramètre  $\sigma_s$  témoigne d'une relative sensibilité aux états de la machine, que ce soit à vide ou en charge.



**Fig. 3. 10.** Variations de l'écart – type de  $I_{s\alpha,\beta}$ . Machine saine (\*). Rotor en défaut (o). Déséquilibre stator ( $\Delta$ )

En conséquence, nous avons calculé deux couples de paramètres ( $m_p, m_q$  et  $\sigma_\alpha, \sigma_\beta$ ) aptes à fournir un renseignement sur la puissance mise en jeu et, deux paramètres ( $\zeta$  et  $\sigma_s$ ) qui serviront plutôt à détecter, voire identifier, les défauts.

Les caractéristiques  $I_{s\beta}(I_{s\alpha})$  (figures 3.11 et 3.12) normalisées par rapport à la valeur efficace de  $I_{s\alpha,\beta}$  (3.12) ont permis de calculer deux autres paramètres.

$$I_{s\alpha,\beta \text{ eff}} = \left[ \frac{1}{N_p} \sum_{k=1}^{N_p} I_{s\alpha,\beta}^2(k) \right]^{1/2} \quad (3.12)$$

La caractéristique  $I_{s\beta}(I_{s\alpha})$  traduit l'évolution du courant stator  $\underline{I}_{s\alpha,\beta}$  dans le plan complexe :

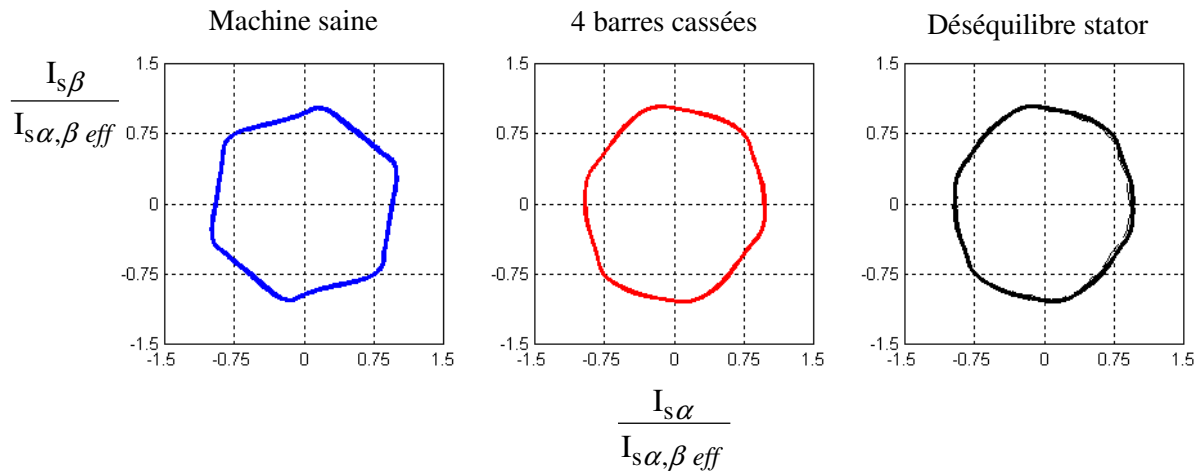
$$\underline{I}_{s\alpha,\beta} = I_{s\alpha} + jI_{s\beta} = I_{s\alpha,\beta} e^{j\varphi} \quad (3.13)$$

Les deux paramètres, notés  $\delta_\alpha$  et  $\delta_\beta$  sont les mesures de la déformation de cette caractéristique à  $\varphi=0$  et à  $\varphi=\frac{\pi}{2}$ , autrement dit dans les axes  $\alpha$  et  $\beta$  :

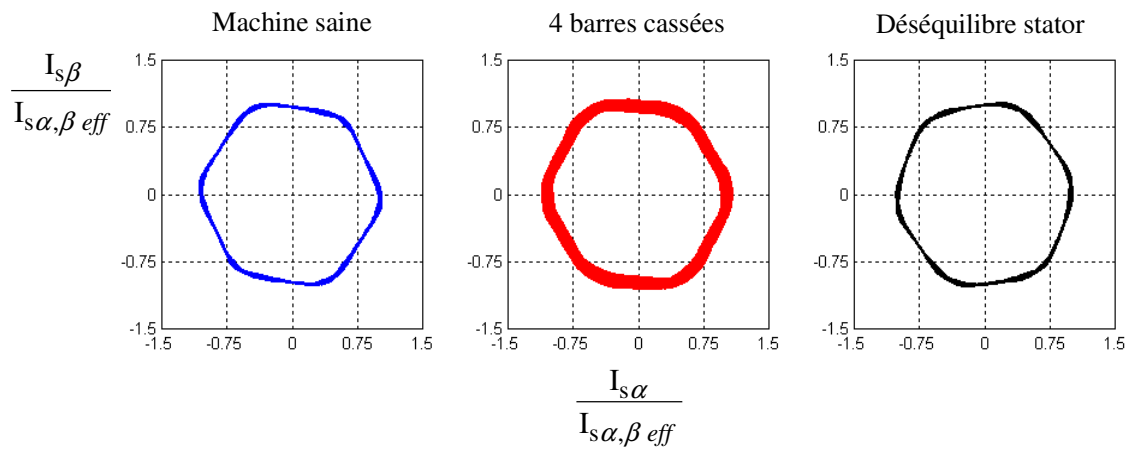


$$\begin{cases} \delta_\alpha = \left| \max(I_{s\alpha,\beta}) - \min(I_{s\alpha,\beta}) \right|_{\varphi=0} \\ \delta_\beta = \left| \max(I_{s\alpha,\beta}) - \min(I_{s\alpha,\beta}) \right|_{\varphi=\frac{\pi}{2}} \end{cases} \quad (3.14)$$

Après analyse des caractéristiques, le premier constat est que la différence est difficile à faire pour les fonctionnements à vide. Par contre, pour le défaut "barres cassées" en charge la caractéristique offre des résultats relativement intéressants.

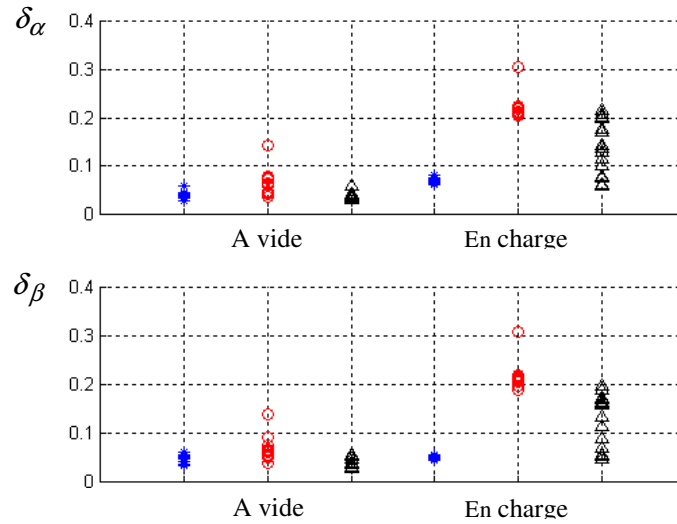


**Fig. 3. 11.** Caractéristique  $I_{s\beta}(I_{s\alpha})$  normalisée pour le fonctionnement à vide



**Fig. 3. 12.** Caractéristique  $I_{s\beta}(I_{s\alpha})$  normalisée pour le fonctionnement en charge

La distinction entre les différents modes de fonctionnement, fournie par l'analyse des paramètres de déformation (figure 3.13) est plutôt moyenne, en tout cas pour le fonctionnement à vide. Pour le fonctionnement en charge, une forte dispersion de la classe représentant le défaut stator est à prévoir si ces paramètres sont sélectionnés.



**Fig. 3. 13.** Déformation de la caractéristique  $I_{s\beta}(I_{s\alpha})$  normalisée pour l'ensemble des acquisitions.

### 3.3.1.2. Données fréquentielles

#### 3.3.1.2.1 Analyse du courant stator

La signature fréquentielle spécifique au défaut "barres cassées" est la forte modulation des harmoniques du courant par la fréquence de glissement. Dans cette analyse fréquentielle, il est préférable de faire une étude globale sur les courants. Il est vrai qu'un défaut rotor est visible sur le spectre des trois courants de ligne. Mais faire une FFT sur chaque courant risque d'être pénalisant en temps de calcul. C'est pour cela que nous avons opté pour l'analyse spectrale de la combinaison des trois courants, traduite par la composante symétrique direct  $\underline{I}_{sd}$ .

$$\underline{I}_{sd} = I_{sa} + \underline{a} \cdot I_{sb} + \underline{a}^2 \cdot I_{sc} \quad (3.15)$$

Où l'opérateur  $\underline{a}$  représente une rotation de  $\frac{2\pi}{3}$  :  $\underline{a} = e^{\frac{j2\pi}{3}}$ . Nous reviendrons par la suite sur les composantes symétriques directes et inverses pour d'autres calculs de paramètres. L'analyse spectrale a été réalisée sur la partie réelle de  $\underline{I}_{sd}$ .

Dans le cas du défaut rotor, nous rappelons (Chapitre 1) que les principales raies pouvant servir de signatures sont localisées aux fréquences suivantes :

- Autour de la fréquence d'alimentation :

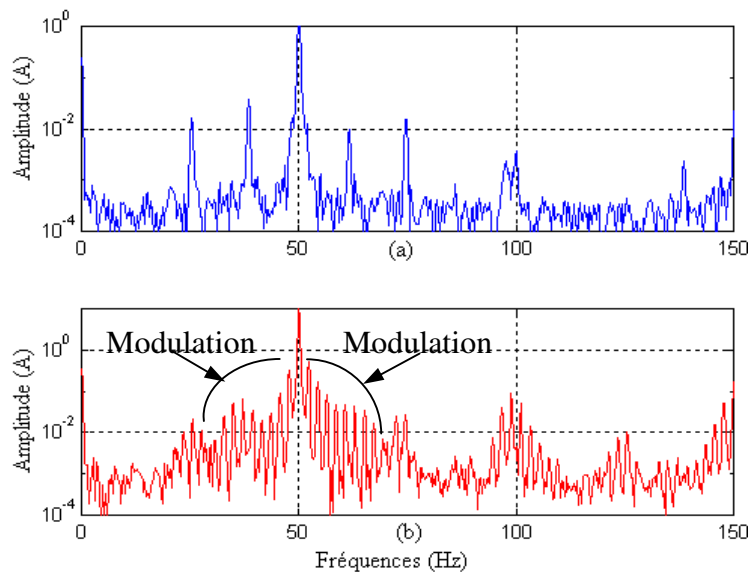
$$f = F_s (1 \pm 2 \cdot n \cdot g) \quad (3.16)$$

- Autour des fréquences principales d'encoches :

$$f = F_s \left[ \frac{Nr(1-g)}{p} \pm 2 \cdot m \right] \pm 2 \cdot k \cdot g \cdot F_s \quad (3.17)$$

Où  $n$ ,  $m$ , et  $k$  sont des entiers.

La figure 3.14, sur laquelle la modulation d'amplitude est visible, corrobore l'expression (3.16).



**Fig. 3. 14.** Spectre du courant  $I_{sd}$ , fonctionnement en charge. Rotor sain (a), rotor en défaut (b)

### 3.3.1.2.2 Amplitudes des raies aux fréquences caractéristiques

La valeur du glissement a permis de calculer les fréquences caractéristiques.

Soit  $n'$  (en tr/mn) la vitesse de rotation du rotor et  $n$  celle du champ tournant, alors le glissement peut également être défini :

$$g = \frac{n - n'}{n} \quad (3.18)$$

Les raies sont situées généralement autour de la fréquence d'alimentation et autour des fréquences principales d'encoches. Les fréquences sont exprimées dans le tableau (3.2).

Expression des fréquences	Description
$f_s(1 \pm 2g)$	Modulation autour du fondamental (soit 4 valeurs)
$f_s(1 \pm 4g)$	
$f_s \pm f_r$	Raies pouvant traduire un éventuel désalignement (2 valeurs)
$3f_s$	Harmonique 3 du fondamental
$N_r f_r \pm f_s \pm 2 \cdot m \cdot g \cdot f_s$	Fréquences d'encoches rotor et modulation avec $m = 0$ ou $1$ (6 valeurs)
$N_r f_r \pm 2f_s \pm 2 \cdot m \cdot g \cdot f_s$	"Harmoniques 2" des fréquences d'encoches rotor et modulation avec $m = 0$ ou $1$ (6 valeurs)
$N_r f_r \pm 3f_s \pm 2 \cdot m \cdot g \cdot f_s$	"Harmoniques 3" des fréquences d'encoches rotor et modulation avec $m = 0$ ou $1$ (6 valeurs)

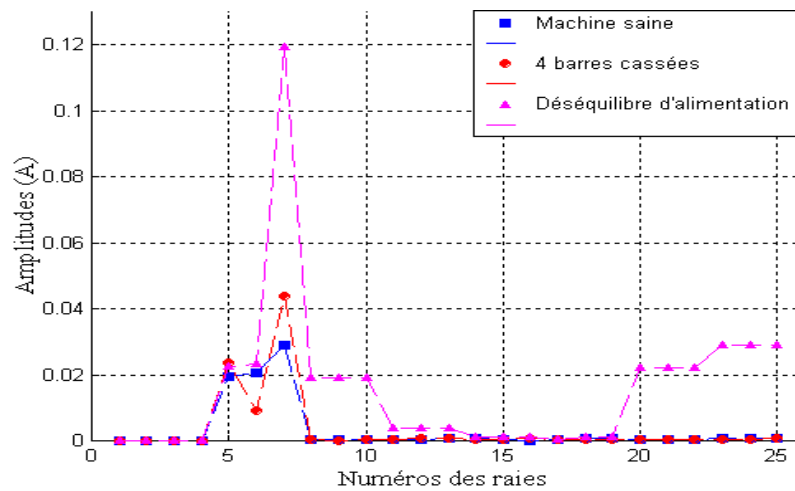
**Tableau. 3. 2.** Expression des fréquences recherchées

Au total, 25 amplitudes de raies sont calculées pour les paramètres fréquentiels. Afin d'analyser la variation de ces paramètres, nous avons effectué, pour chacune des amplitudes de raies, une moyenne sur l'ensemble des 15 acquisitions dans chaque cas.

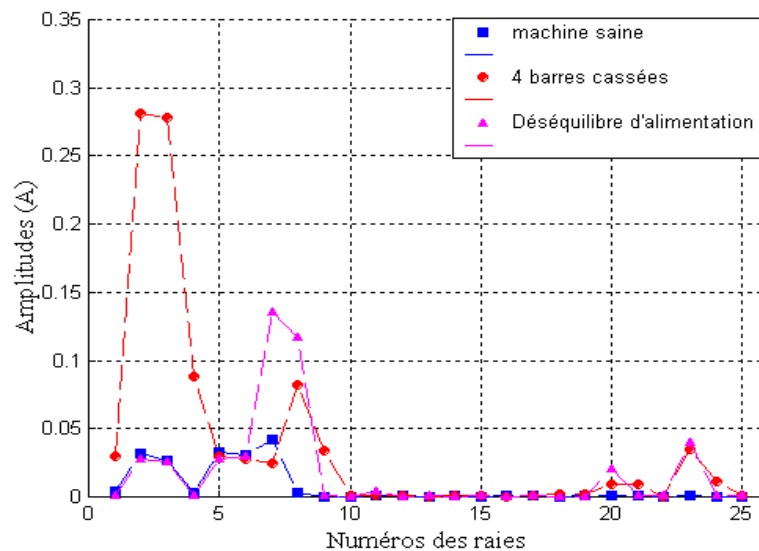
Les valeurs correspondantes sont reprises sur les figures 3.15 et 3.16 dans l'ordre suivant :

- les quatre premières valeurs sont les moyennes des amplitudes des raies de modulation autour du fondamental, rangées dans l'ordre croissant d'apparition des fréquences.
- les valeurs 5, 6, et 7 correspondent respectivement aux fréquences suivantes :  $f_s - f_r$ ,  $f_s + f_r$  et  $3 \cdot f_s$
- Enfin les valeurs 8 à 25 correspondent aux fréquences d'encoches (tableau 3.3) rangées dans l'ordre croissant d'apparition des fréquences.

Par ailleurs, à vide, les raies de modulation à  $f_s(1 \pm 2 \cdot n \cdot g)$   $n = 1, 2$  ont volontairement été annulées car elles ne sont pas réalistes (confusion avec le fondamental).



**Fig. 3.15.** Variation des raies correspondant aux fréquences recherchées. Fonctionnement à vide.



**Fig. 3.16.** Variation des raies correspondant aux fréquences recherchées. Fonctionnement en charge.

Le premier constat est une réaction importante de la raie à  $3 \cdot f_s$  (n° 7) au défaut de déséquilibre à vide et en charge. A vide et autour de  $N_r f_r \pm f_s$  et de  $N_r f_r \pm 3f_s$ , nous remarquons aussi une plus grande variation pour le déséquilibre que pour les autres défauts.

Ces résultats mettent en évidence la possibilité de différencier le défaut stator du défaut rotor en recherchant les raies aux multiples de l'harmonique trois du fondamental et autour des fréquences principales d'encoches. Les raies 5 et 6 n'étant jamais nulles et ayant les mêmes valeurs quel que soit le mode de fonctionnement, nous concluons que l'arbre de la machine n'est probablement pas parfaitement aligné.

L'analyse des composantes à  $f_s(1 \pm 2 \cdot n \cdot g)$  en charge fournit, comme prévu, un résultat net du point de vue de l'identification du défaut rotor. D'autre part la distinction entre les différents modes est une fois de plus possible si les amplitudes des raies aux fréquences  $3f_s$  et  $N_r f_r \pm f_s(1 \pm 2 \cdot m \cdot g)$  sont analysées.

Afin de compléter la recherche de paramètres par analyse spectrale, un calcul énergétique a été réalisé.

### 3.3.1.2.3 Calcul énergétique

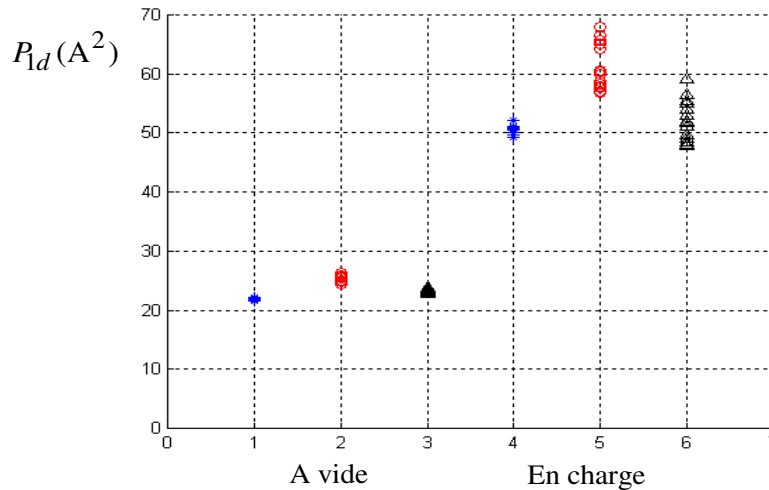
Les signaux étant sinusoïdaux (donc à énergie infinie), leurs puissances moyennes sont évaluées. Cette analyse vise à mettre en évidence, d'une manière globale, les changements dans le spectre du courant. Désignons par  $I_d(f)$  la transformée de Fourier de  $\text{Re}[I_{sd}]$  dont la densité spectrale de puissance est tout simplement :

$$S_d(f) = I_d(f) \cdot \overline{I_d(f)} = |I_d(f)|^2 \tag{3.19}$$

La puissance moyenne du signal vaut donc :

$$P_{1d} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} S_d(f) df \tag{3.20}$$

Ce paramètre traduit une relative sensibilité aussi bien à vide qu'en charge.



**Fig. 3. 17.** Variations de la puissance moyenne du courant. Machine saine (\*). Rotor en défaut (o). Déséquilibre stator ( $\Delta$ )

### 3.3.1.3. Composantes symétriques directes et inverses

L'expression de la composante symétrique directe des courants est donnée par l'équation (3.15).

Mais d'une manière générale, tout système de trois grandeurs de même nature et de même fréquence est la superposition de trois systèmes équilibrés de même fréquence : un système triphasé direct, un système triphasé inverse et un système homopolaire.

En pratique, cela signifie que le système des trois tensions ou des trois courants, peut être représenté par les composantes  $\underline{X}_d$ ,  $\underline{X}_i$  et  $\underline{X}_0$  (tensions ou courants) telles que :

- $\underline{X}_d$  représente la composante symétrique directe
- $\underline{X}_i$  représente la composante symétrique inverse
- $\underline{X}_0$  représente la composante homopolaire

Les équations donnant les composantes symétriques des tensions et des courants (transformation de Fortescue) sont les suivantes :

$$\begin{cases} \underline{X}_d = \frac{1}{3} \cdot (\underline{X}_{sa} + \underline{a} \cdot \underline{X}_{sb} + \underline{a}^2 \cdot \underline{X}_{sc}) \\ \underline{X}_i = \frac{1}{3} \cdot (\underline{X}_{sa} + \underline{a}^2 \cdot \underline{X}_{sb} + \underline{a} \cdot \underline{X}_{sc}) \\ \underline{X}_0 = \frac{1}{3} \cdot (\underline{X}_{sa} + \underline{X}_{sb} + \underline{X}_{sc}) \end{cases} \quad (3.21)$$

Il s'agit par cette approche d'avoir une estimation de l'impédance stator par le biais des impédances directes et inverses.

Des travaux [Cabanas, 97] et [Assaf, 01] ont montré l'importance du calcul des impédances directes et inverses pour la détection de spires en court – circuit ou du déséquilibre d'alimentation.

Les deux paragraphes suivants présentent le calcul simple de ces impédances.

#### 3.3.1.3.1 Impédance directe

Prenons les expressions de  $\underline{V}_d$  et  $\underline{I}_d$  (3.24), et les spectres de leurs modules que nous désignerons simplement par  $V_{sd}(f)$  et  $I_{sd}(f)$ . L'impédance directe  $Z_d$  est le rapport entre le fondamental de  $V_{sd}(f)$  et celui de  $I_{sd}(f)$  :

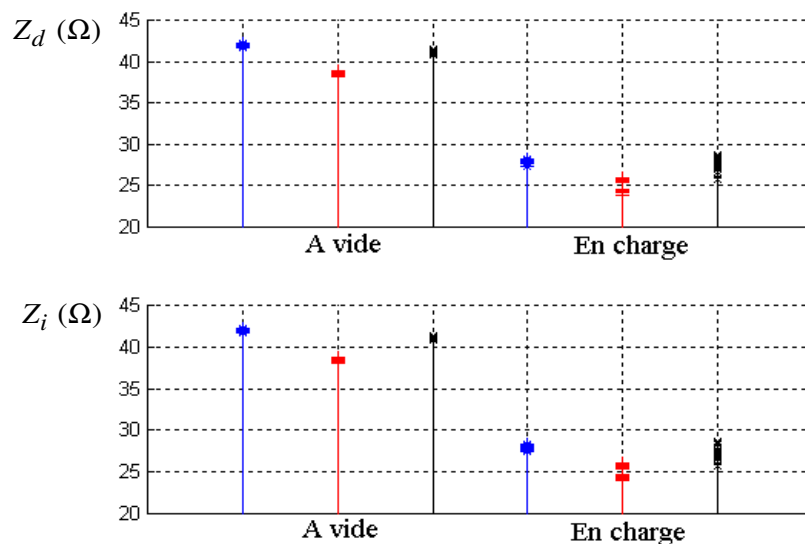
$$Z_d = \left. \frac{V_{sd}}{I_{sd}} \right|_{f=50\text{Hz}} \quad (3.22)$$

### 3.3.1.3.2 Impédance inverse

Il en est de même pour l'impédance inverse  $Z_i$  qui, cette fois, est le rapport entre  $V_{si}(f)$  et  $I_{si}(f)$  à la fréquence du fondamental :

$$Z_i = \left. \frac{V_{si}}{I_{si}} \right|_{f=50\text{Hz}} \quad (3.23)$$

La figure 3.18 montre que, pour chacun des modes de fonctionnement présentés, les deux valeurs d'impédance varient dans les mêmes proportions. L'information fournie par ces deux paramètres risque d'être redondante.



**Fig. 3.18.** Variation de l'impédance directe et de l'impédance inverse. Machine saine (\*). Rotor en défaut (+). Déséquilibre stator (x)

Dans le vecteur forme final, l'un de ces paramètres aidera certainement à distinguer le défaut stator du défaut rotor. Du fait du faible niveau de déséquilibre d'alimentation sur une phase (environ 10 %), la distinction entre ce défaut et le fonctionnement sain n'est pas nette. Mais cette limitation pourrait être compensée par les paramètres temporels et fréquentiels.

Toutefois, la matrice de variance – covariance et la matrice de dispersion intra – classe ne seront pas inversibles si ces deux paramètres sont conservés. En effet,  $Z_d$  et  $Z_i$  sont égaux



dans tous les cas de fonctionnement. C'est donc l'impédance directe qui est utilisée dans le vecteur forme. Ce choix est purement arbitraire.

### 3.3.2 Choix du vecteur forme

En plus du choix de ne garder que l'impédance directe, il est possible de réduire encore la dimension de l'espace de départ en éliminant d'autres paramètres. En effet, les amplitudes aux fréquences de modulation  $f_s(1 \pm 2 \cdot n \cdot g)$  ( $n = 1, 2$ ) ne sont pas réalistes pour les fonctionnements à très faible charge. Or, nous souhaitons réaliser un diagnostic également pour ce type de fonctionnement. De ce fait, ces paramètres ne seront pas conservés dans le vecteur forme. Toutefois l'analyse de ces raies pourrait être utilisée par la suite pour aider à l'identification des observations rejetées.

La signature de la machine est définie par un vecteur de trente et un paramètres. Nous présentons un panorama complet de l'ensemble de ces paramètres sur le tableau 3.19.

Quelques-uns de ces paramètres risquent d'être non pertinents ou redondants selon l'ensemble d'apprentissage à définir. Le dimensionnement de l'espace est donc revu à la baisse grâce à la sélection de paramètres.

Il est toujours possible, pour l'utilisateur, de définir un vecteur forme en se basant sur les analyses de sensibilité des paramètres. L'approche par expertise humaine peut donner de très bons résultats comme c'est le cas pour les études [Casimir, 03 a] et [Casimir, 03 b]. Toutefois, la sélection de caractères rend la recherche de l'espace de représentation plus rapide et plus globale.

Des algorithmes de sélection ont été présentés au cours du chapitre précédent. Deux d'entre eux seront utilisés pour définir la signature caractérisant l'ensemble d'apprentissage. Il s'agit de la méthode calculant le rapport de Fisher sur chaque paramètre et de la méthode dite "Sequential Backward Selection" (SBS).

Les algorithmes de sélection sont chargés de rechercher les paramètres les plus pertinents conformément à la dimension de l'espace fixée par l'utilisateur.

Néanmoins, le vecteur forme doit permettre de distinguer clairement les différentes classes. De ce fait, la visualisation des classes dans le plan factoriel, obtenu par analyse en composantes principales, permet d'avoir une première idée du caractère discriminant. La sélection sera validée par une procédure de test.

$\alpha_i$	Dénomination		$\alpha_i$	Dénomination	
1	$\zeta$	Dispersion des points de la caractéristique Q(P)	17	$N_r f_r + 2 f_s$	
2	$m_p$	Puissance active moyenne	18	$N_r f_r + 2 f_s + 2 g f_s$	
3	$m_q$	Puissance réactive moyenne	19	$N_r f_r - 3 f_s - 2 g f_s$	
4	$f_s - f_r$	Amplitudes aux fréquences $f_s \pm f_r$	20	$N_r f_r - 3 f_s$	Amplitudes autour de $N_r f_r \pm 3 f_s$
5	$f_s + f_r$		21	$N_r f_r - 3 f_s + 2 g f_s$	
6	$3 f_s$		Harmonique 3 du fondamental	22	
7	$N_r f_r - f_s(1 + 2g)$	Amplitudes autour de $N_r f_r \pm f_s$	23	$N_r f_r + 3 f_s$	
8	$N_r f_r - f_s$		24	$N_r f_r + 3 f_s + 2 g f_s$	
9	$N_r f_r - f_s(1 - 2g)$		25	$P_{1d}$	
10	$N_r f_r + f_s(1 - 2g)$		26	$\delta_\alpha$	
11	$N_r f_r + f_s$		27	$\delta_\beta$	
12	$N_r f_r + f_s(1 + 2g)$		28	$\sigma_\alpha$	
13	$N_r f_r - 2 f_s - 2 g f_s$	Amplitudes autour de $N_r f_r \pm 2 f_s$	29	$\sigma_\beta$	
14	$N_r f_r - 2 f_s$		30	$\sigma_s$	
15	$N_r f_r - 2 f_s + 2 g f_s$		31	$Z_d$	
16	$N_r f_r + 2 f_s - 2 g f_s$				

**Tableau. 3. 3.** Panorama des paramètres calculés

Les paramètres de départ ne sont pas homogènes. La sélection des paramètres ainsi que l'analyse en composantes principales risquent d'être biaisées si l'ensemble d'apprentissage n'est pas normalisé. Nous n'utiliserons que des données normalisées par la suite.

Les deux paragraphes suivants présentent la sélection des paramètres en fonction de la dimension désirée et de la méthode utilisée. Afin d'illustrer les résultats, les algorithmes de sélection sont appliqués, dans un premier temps, aux données recueillies sur les modes de fonctionnement en alimentation réseau suivants :

- Fonctionnement machine saine, à vide et en charge
- 4 barres cassées, à vide et en charge
- Déséquilibre au stator, à vide et en charge

### 3.3.2.1. Sélection par le critère de Fisher

Le mode de sélection est relativement simple : à l'aide de l'ensemble d'apprentissage, le rapport de Fisher (équation 2.20 ch. II) est calculé pour chaque paramètre. Ce calcul est une évaluation de la capacité de chaque paramètre à différencier les classes. Un paramètre  $\alpha$  aura un pouvoir de séparation d'autant meilleur que le critère de Fisher qui lui est associé est grand.

Les paramètres retenus sont ceux qui correspondent aux  $d'$  plus fortes valeurs du critère.

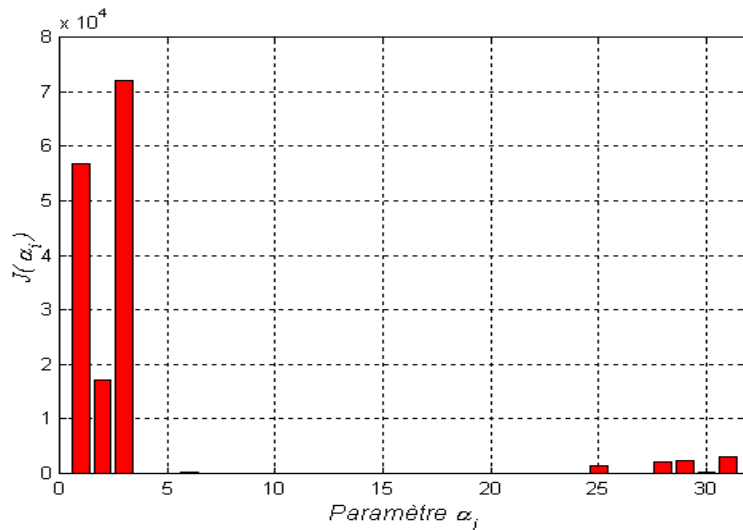
Soit  $\{\alpha_i\}$  ( $i = 1, \dots, d'$ ) l'ensemble des paramètres  $\alpha_i$  rangés par ordre décroissant de leurs valeurs de critère. Le vecteur forme, qui caractérisera l'ensemble d'apprentissage, est :

$$\mathbf{x} = [x_1, x_2, \dots, x_{d'}]^t = [\alpha_1, \alpha_2, \dots, \alpha_{d'}]^t \quad (3.24)$$

Avec :

$$J(\alpha_1) > J(\alpha_2) > \dots > J(\alpha_{d'}) \quad (3.25)$$

Afin d'illustrer la méthode de sélection, l'histogramme suivant montre la variation du critère pour les 31 paramètres calculés.



**Fig. 3. 19.** Variation du rapport de Fisher pour chaque paramètre

Le critère de Fisher dit qu'il y a 7 paramètres qui, pris séparément, ont un pouvoir discriminant. La figure 3.19 montre clairement que les paramètres  $\alpha_3$ ,  $\alpha_1$  et  $\alpha_2$  ont les pouvoirs discriminants les plus élevés. Mais ce résultat ne signifie pas pour autant qu'il faille limiter la taille du vecteur forme à trois. En effet, celui – ci doit être défini sur la base d'un apport collectif des paramètres discriminants.

Donc d'après le résultat de la figure 3.19, la dimension du vecteur forme peut aller de 2 à 7.

Selon la procédure présentée sur la figure 3.4 le meilleur vecteur forme, parmi les 6 solutions possibles, est celui qui donnera les meilleurs performances au sens du nombre de d'observations de test mal classées.

Le tableau 3.4 donne le nombre d'observations mal classées dans toutes les configurations de vecteur forme possibles. Cette analyse a été réalisée à l'aide de la règle des "5 – ppv".

Paramètres du vecteur forme	Nombre de mal classés
$[\alpha_3, \alpha_1]$	4
$[\alpha_3, \alpha_1, \alpha_2]$	1
$[\alpha_3, \alpha_1, \alpha_2, \alpha_{31}]$	0
$[\alpha_3, \alpha_1, \alpha_2, \alpha_{31}, \alpha_{29}]$	0
$[\alpha_3, \alpha_1, \alpha_2, \alpha_{31}, \alpha_{29}, \alpha_{28}]$	0
$[\alpha_3, \alpha_1, \alpha_2, \alpha_{31}, \alpha_{29}, \alpha_{28}, \alpha_{25}]$	0

**Tableau. 3. 4.** Resultats de test pour définir le vecteur forme à l'aide du critère de Fisher

Par conséquent, nous pouvons prendre pour vecteur forme :  $\mathbf{x} = [\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_{31}]^t$ .

Les paramètres correspondent respectivement à :

- La dispersion des points de la caractéristique normalisée  $Q'(P)$  :  $\zeta$
- Les moyennes des puissances actives et réactives :  $m_p, m_q$
- L'impédance directe :  $Z_d$ .

### 3.3.2.2. Sélection par la méthode SBS

Contrairement à la méthode précédente, l'algorithme SBS utilise le calcul d'un critère sur des ensembles de paramètres.

Le critère utilisé est le rapport entre le déterminant de la matrice de variance – covariance  $\mathbf{S}$  et le déterminant de la matrice de variance intra – classe  $\mathbf{S}_w$  :

$$J = \frac{\det(\mathbf{S})}{\det(\mathbf{S}_w)} \quad (3.26)$$

L'espace de départ est l'ensemble des paramètres calculés. A chaque étape de la procédure, le sous – espace qui maximise le critère est pris en considération. Il ne s'agit pas d'éliminer le paramètre le moins pertinent, mais de conserver le sous – ensemble qui caractérise le mieux l'espace de représentation en accord avec la dimension désirée.

La difficulté est de savoir à quelle itération arrêter la procédure. Le critère  $J(E_k)$  (équation 2.25, chapitre 2) est un bon critère d'arrêt. La figure 3.20 illustre la réalisation de la sélection des sous – ensemble des paramètres les plus pertinents. Et la figure 3.21 donne, à chaque étape de la sélection, la valeur du critère  $J(E_k)$ .

Sur cette courbe, seule la zone de plus grande pente comprise entre les itérations 25 et 29 est intéressante. La dimension du vecteur forme peut donc aller de 5 à 2. Après avoir appliqué la procédure de la figure 3.4, nous avons obtenu les résultats de classification du tableau 3.5.

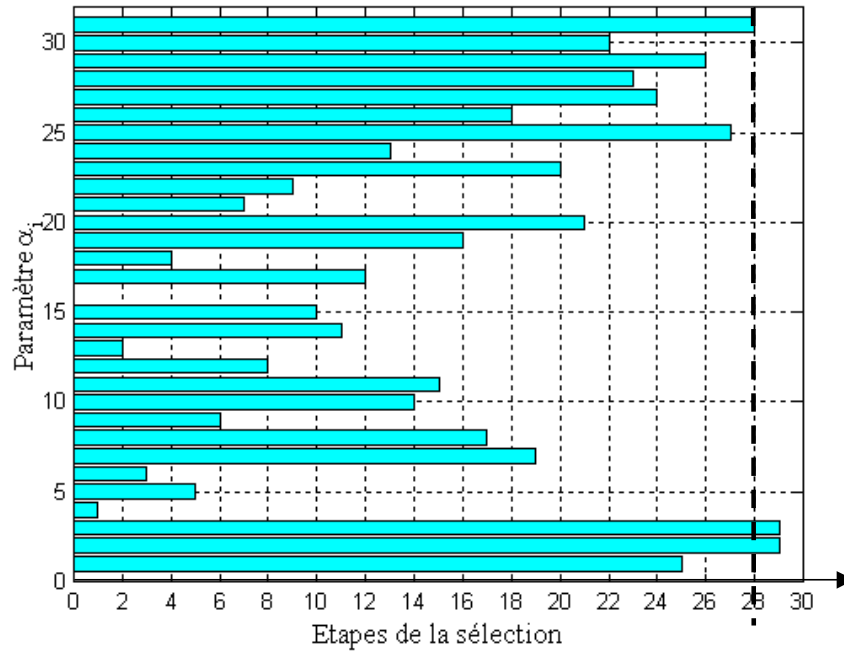


Fig. 3. 20. Procédure de sélection des paramètres par la méthode SBS

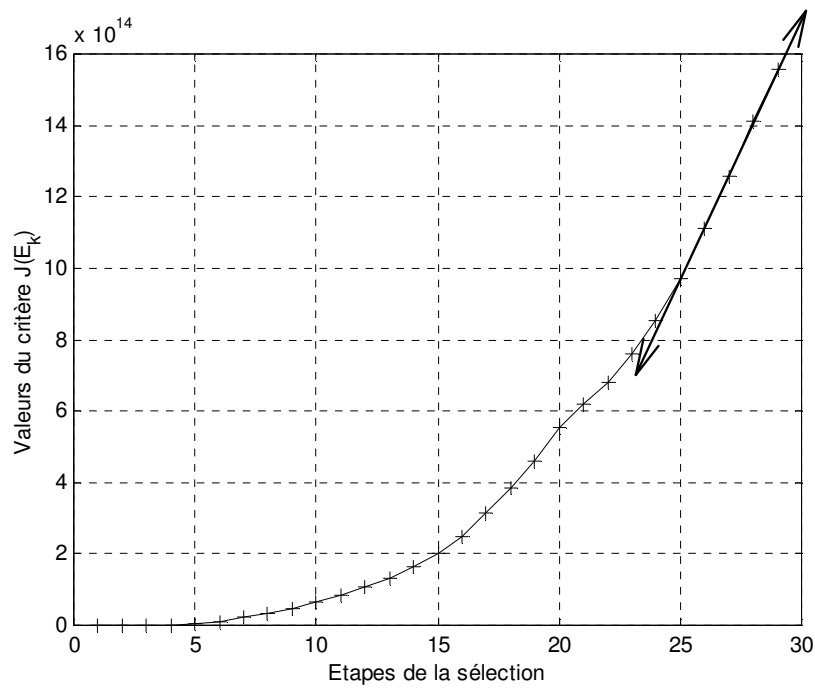


Fig. 3. 21. Valeur du critère  $J(E_k)$  à chaque étape de la sélection SBS.

Paramètres du vecteur forme	Nombre de mal classés
$[\alpha_2, \alpha_3]$	4
$[\alpha_2, \alpha_3, \alpha_{31}]$	0
$[\alpha_2, \alpha_3, \alpha_{31}, \alpha_{25}]$	0
$[\alpha_2, \alpha_3, \alpha_{31}, \alpha_{25}, \alpha_{29}]$	0

**Tableau. 3. 5.** *Resultats de test pour définir le vecteur forme à l'aide de l'algorithme SBS*

Par conséquent, la procédure de sélection est arrêtée lorsque  $d' = 3$  (c'est à dire à la 28<sup>ième</sup> étape), le sous – ensemble de paramètres constituant le vecteur forme est composé des paramètres  $\alpha_2$ ,  $\alpha_3$ , et  $\alpha_{31}$ , c'est à dire :

- les puissances actives et réactives moyennes  $m_p$  et  $m_q$
- et l'impédance direct  $Z_d$

Que l'on utilise le critère de Fisher ou la méthode SBS, les paramètres du vecteur forme sont presque toujours les mêmes. Sur la base de cet exemple, le rapport de Fisher donne un vecteur forme incluant forcément le paramètre  $\alpha_1$ . Ce paramètre ne s'avère pas pertinent dans le cadre d'un apport collectif pris en compte par la méthode SBS. L'autre avantage de la méthode SBS est que l'analyse factorielle ne sera pas nécessaire pour visualiser les données car le vecteur forme est de dimension 3.

Les résultats de la méthodes de sélection de caractères ont été appliqués aux données issues des modes de fonctionnement à vide et pour une charge d'environ 3 kW. L'analyse en composantes principales permettra de visualiser l'ensemble des classes.

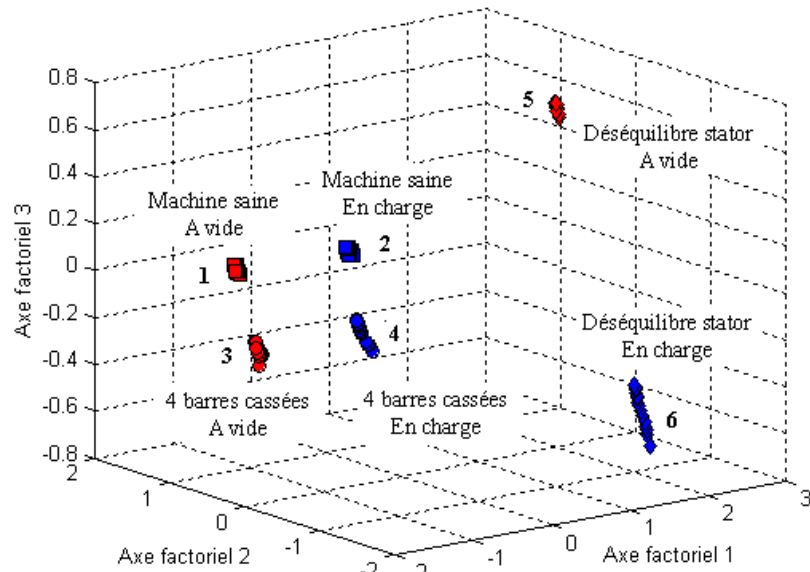
Pour simplifier les notations, la dimension de l'espace sera dorénavant considérée comme étant égale à  $d$  (nombre de paramètres sélectionnés), tant qu'il n'y aura aucune ambiguïté.

### 3.3.3 Visualisation par analyse en composantes principales

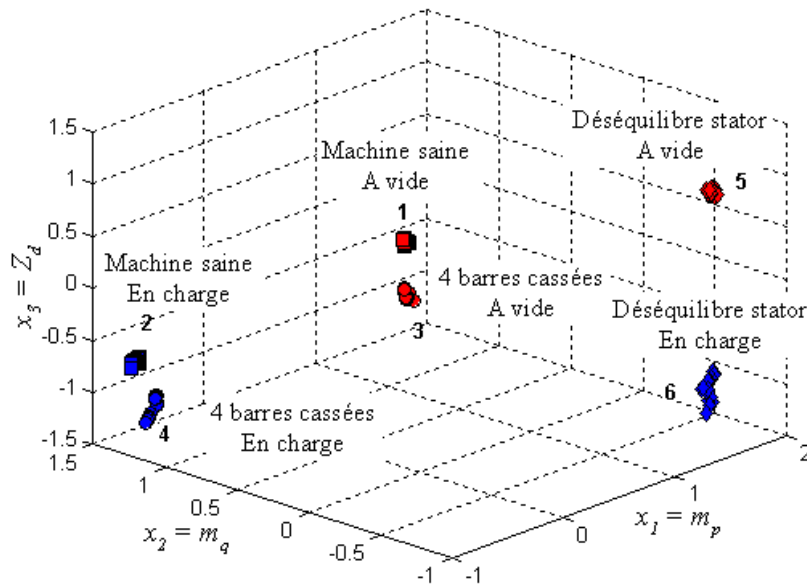
La position des données dans l'espace peut être visualisée après projection de ces dernières dans le plan engendré par les axes 1 et 2 ou 1 et 3. Une visualisation plus globale en dimension 3 est bien entendue possible, surtout si l'inertie expliquée (3.27) est proche de 1.

$$\Gamma = \frac{\sum_{i=1}^3 \lambda_i}{\sum_{j=1}^d \lambda_j} \tag{3.27}$$

Les figures 3.22 et 3.23 montrent la position et la forme des classes dans un espace de dimension trois, à partir des vecteur formes obtenus par les deux algorithmes de sélection.



**Fig. 3. 22.** Position des classes dans le sous – espace de dimension 3. Les paramètres sont sélectionnés avec le critère de Fisher.



**Fig. 3. 23.** Position des classes en dimension 3. Les paramètres sont sélectionnés avec l’algorithme SBS.



Le vecteur forme défini dans les deux cas s'avère pertinent. En effet les classes sont parfaitement distinctes. Ce constat prouve que le caractère non optimal de ces méthodes de sélection n'empêche pas d'obtenir un bon résultat.

L'analyse de la contribution des valeurs propres permet de quantifier le pourcentage d'inertie après projection dans le sous – espace permettant de visualiser les données définies par le critère de Fisher. Ces données sont disponibles sur le tableau suivant.

La valeur du pourcentage d'inertie cumulé par les trois premières valeurs propres montre que toute l'information a été conservée lors de la projection.

Vecteur forme par le rapport de Fisher			
$\mathbf{x} = [\zeta, m_p, m_q, Z_d]^t$			
	Valeur propre	Pourcentage	Pourcentage cumulé
	$\lambda_1$	2,45	61,37
	$\lambda_2$	1.39	34,71
	$\lambda_3$	0.15	3,91
	$\lambda_4$	0.0001	0.003
			100

**Tableau. 3. 6.** Contribution des valeurs propres à la projection des données définies par le critère de Fisher.

D'autre part, les composantes des vecteurs propres sont des combinaisons linéaires des paramètres du vecteur forme. En d'autres termes, les vecteurs propres ont un double rôle dans la projection. Ils constituent la base du sous – espace et renseignent sur la contribution des paramètres (donc des composantes des vecteurs  $\mathbf{x}$  de l'ensemble d'apprentissage) à la construction des axes.

L'ensemble d'apprentissage étant normalisé, les composantes des vecteurs propres seront comprises entre  $-1$  et  $1$ . L'attention est portée sur les trois premiers vecteurs propres car les données sont projetées dans un espace de dimension 3. Les composantes des vecteurs propres dans la matrice suivante montrent que chaque paramètre contribue plus ou moins, en valeur absolue, à la description de l'ensemble d'apprentissage sur chacun des axes.

$$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \mathbf{v}_4 \\ 0,61 & -0,22 & 0,10 & 0,75 \\ -0,53 & -0,40 & 0,71 & 0,21 \\ -0,55 & -0,41 & -0,36 & 0,63 \\ 0,19 & 0,78 & 0,59 & 0,0004 \end{bmatrix} \begin{matrix} \leftarrow \zeta \\ \leftarrow m_p \\ \leftarrow m_q \\ \leftarrow Z_d \end{matrix} \quad (3.28)$$

Comme ce fut précisé auparavant, la projection des données dans le sous – espace permet de visualiser les classes, ce qui confirme visuellement le pouvoir discriminant des paramètres.

### 3.3.4 Conclusion

Le prétraitement des signaux et la recherche du vecteur forme constituent une étape très importante de la phase d'apprentissage du système de diagnostic par reconnaissance des formes. Dans cette partie, nous avons montré le processus de construction du vecteur forme et l'efficacité des méthodes de réduction de sa dimension.

Cette étude sur la recherche de signature, réalisée avec succès, prouve la faisabilité du diagnostic par RDF. En effet, d'une manière générale les résultats du diagnostic sur un processus dépendent de la pertinence des analyses effectuées (signatures).

### 3.4 Procédure de décision

Les acquisitions effectuées à vide et pour une charge de 3kW sur le moteur sain, le moteur avec quatre barres cassées et le déséquilibre d'alimentation ont permis de valider les algorithmes de reconnaissance des formes.

Le diagnostic de la machine asynchrone va être établi à partir d'un ensemble plus riche et diversifié de modes de fonctionnement en alimentation réseau, à savoir :

- la machine saine et avec trois barres cassées pour les niveaux de charge suivants : 0%, 25%, 50%, 75%, et 100%.
- le déséquilibre d'alimentation et une barre partiellement cassée pour 0%, 25%, 50%, et 75% de charge.

Il s'agit maintenant de mettre en place la procédure de décision complète. Comme précisé dans l'introduction de ce chapitre, nous utilisons la seconde campagne d'acquisition (voir tableau 3.1). Cette base de connaissance est scindée en trois ensembles : un ensemble d'apprentissage, un ensemble de test et un ensemble de données pour la phase d'exploitation.

L'ensemble d'apprentissage  $\mathbf{X}_a$  est constitué de 120 échantillons. Ce sont les ensembles de 10 observations parmi 15 recueillies sur :

- la machine saine avec des niveaux de charge de 0 %, 25 %, 50 % et 75 %
- la machine avec trois barres cassées, avec des niveaux de charge de 0 %, 25 %, 50 % et 75 %
- le déséquilibre d'alimentation statorique respectivement pour 0 %, 25 %, 50 % et 75 % de charge.

L'ensemble de test  $\mathbf{X}_t$  est constitué des 5 mesures non utilisées pour l'apprentissage de :

- la machine saine.
- le défaut rotor.
- le déséquilibre d'alimentation.

L'ensemble de test est donc constitué de 60 observations.

Enfin, l'ensemble de données pour la phase d'exploitation  $\mathbf{X}_u$  contient des modes de

fonctionnement qui n'ont servi ni à l'apprentissage, ni au test.  $\mathbf{X}_u$  est constitué de :

- 15 mesures machine saine, à 100 % de charge
- 15 mesures machine avec un rotor ayant 3 barres cassées, à 100 % de charge
- 15 mesures avec un rotor ayant une barre partiellement cassée pour 0 %, 25 %, 50 % et 75 % de charge.

L'ensemble  $\mathbf{X}_u$  contient 90 échantillons.

La règle des  $k - ppv$  puis le calcul direct des frontières seront utilisés pour la procédure de décision. Nous montrerons l'importance de l'application des règles de rejets d'ambiguïté et de distance pour diminuer le risque de mauvaise classification.

### 3.4.1 Choix du vecteur forme

Il est primordial de définir un vecteur forme caractérisant au mieux l'ensemble d'apprentissage.

Nous ne reprendrons pas ici tout le travail de test réalisé pour chaque jeu de paramètres fourni par les algorithmes de sélection. Seules les signatures qui ont conduit aux meilleurs résultats de classification, en phase de test, sont présentées et utilisées.

Les signatures obtenues au moyen du rapport de Fisher puis par la méthode SBS sont les suivantes :

- $\mathbf{x} = [m_p, m_q, P_{1d}, \sigma_\alpha, \sigma_\beta, Z_d]^t$  sélectionnée à l'aide du rapport de Fisher
- $\mathbf{x} = [m_p, m_q, \sigma_\alpha, \sigma_\beta, \sigma_s, Z_d]^t$  sélectionnée avec l'algorithme SBS

Ces vecteurs formes seront désormais nommés  $VF_1$  pour le rapport de Fisher et  $VF_2$  pour l'algorithme SBS.

Selon le vecteur forme, les figures 3.24 et 3.25 illustrent la projection, par ACP, de l'ensemble d'apprentissage dans un sous – espace de dimension 3.

La désignation des classes est la suivante :

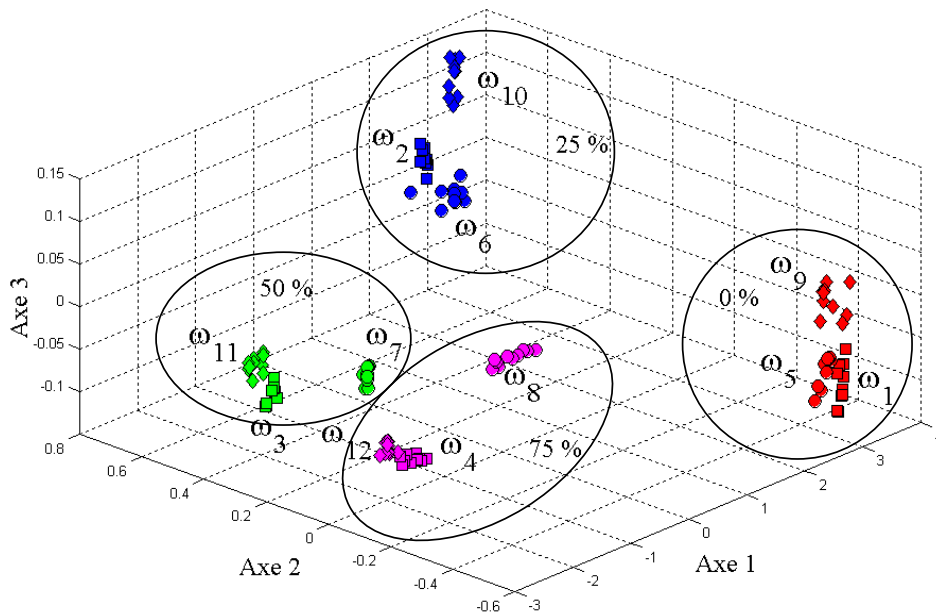
- Les classes  $\omega_1, \omega_2, \omega_3$  et  $\omega_4$  représentent la machine saine de 0% à 75 % de charge respectivement.

- Les classes  $\omega_5, \omega_6, \omega_7$  et  $\omega_8$  représentent la machine avec trois barres cassées de 0% à 75 % de charge respectivement.
- Les classes  $\omega_9, \omega_{10}, \omega_{11}$  et  $\omega_{12}$  représentent le déséquilibre d'alimentation de 0% à 75 % de charge respectivement.

Le pourcentage d'information conservé est de 99,8 % en dimension trois, lors de la projection de l'ensemble (figure 3.24) défini par  $VF_1$ . Les trois premières valeurs – propres  $\lambda_i$  justifient ce pourcentage :

$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	$\lambda_6$
5,75	0,22	$9,6 \cdot 10^{-3}$	$5,3 \cdot 10^{-3}$	$1,1 \cdot 10^{-3}$	$2 \cdot 10^{-5}$

(3.29)



**Fig. 3. 24.** Ensemble d'apprentissage défini par le vecteur forme  $VF_1$ . Machine saine "■", trois barres cassées "●", et déséquilibre stator "◆".

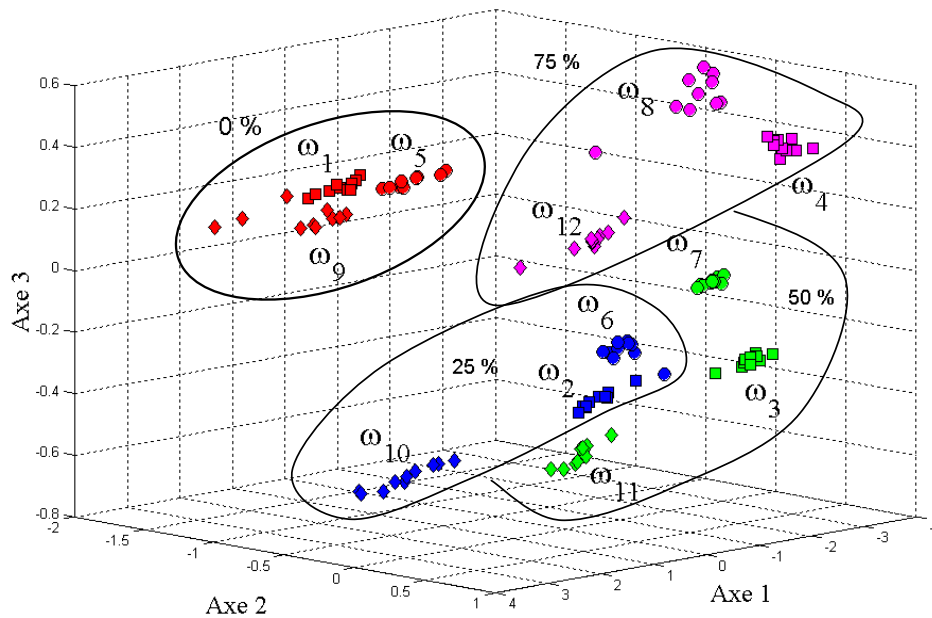
Ce fort taux d'information conservée est confirmé par les composantes des trois premiers vecteurs – propres qui constituent la base du sous – espace (3.30).

En particulier, les valeurs des composantes du premier vecteur – propre montrent les contributions quasi identiques de tous les paramètres sur l'axe 1. Ce dernier porte environ 96 % de l'information, si l'on tient compte de la valeur de  $\lambda_1$ .

$$\begin{aligned}
 v_1 &= [-\mathbf{0,39} \quad \mathbf{0,412} \quad -\mathbf{0,405} \quad -\mathbf{0,412} \quad -\mathbf{0,413} \quad \mathbf{0,415}]^t \\
 v_2 &= [\mathbf{0,739} \quad -\mathbf{0,253} \quad -\mathbf{0,467} \quad -\mathbf{0,298} \quad -\mathbf{0,273} \quad -\mathbf{0,081}]^t \\
 v_3 &= [\mathbf{0,544} \quad \mathbf{0,693} \quad \mathbf{0,356} \quad \mathbf{0,065} \quad \mathbf{0,062} \quad \mathbf{0,296}]^t
 \end{aligned}
 \tag{3.30}$$

Cependant, certaines classes se chevauchent ce qui implique un risque possible d'erreur pour les nouvelles observations.

La situation est différente (figure 3.25) pour les classes définies par  $VF_2$  (le vecteur forme sélectionné par l'algorithme SBS). En effet les classes sont relativement bien séparées et se chevauchent beaucoup moins.



**Fig. 3. 25.** Ensemble d'apprentissage défini par le vecteur forme  $VF_2$ . Machine saine "■", trois barres cassées "●", et déséquilibre stator "◆".

La qualité de la représentation est quasiment identique au cas précédent. Le premier axe conserve moins d'information (88,41 %). Mais la projection en dimension trois est aussi fidèle à la réalité (99,8 %) :

$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	$\lambda_6$
5.3046	0.5401	0.1438	0.0078	0.0034	0.0002

(3.31)

Les trois vecteurs – propres correspondants aux trois premières valeurs – propres sont :

$$\begin{array}{rcccccc}
 & m_p & m_q & \sigma_\alpha & \sigma_\beta & \sigma_s & Z_d \\
 v_1 & = & [-\mathbf{0,405} & -\mathbf{0,428} & -\mathbf{0,426} & -\mathbf{0,429} & -\mathbf{0,319} & \mathbf{0,43}]^t \\
 v_2 & = & [ \mathbf{0,312} & -0,188 & 0,036 & -0,005 & -\mathbf{0,915} & -0,167]^t \\
 v_3 & = & [-\mathbf{0,716} & 0,131 & \mathbf{0,486} & \mathbf{0,411} & -0,238 & -0,09]^t
 \end{array} \tag{3.32}$$

Mis à part le premier axe où tous les paramètres interviennent en proportions égales, ce sont surtout les ensembles  $\{m_p, \sigma_s\}$  et  $\{m_p, \sigma_\alpha, \sigma_\beta\}$  qui interviennent le plus dans la définition des vecteurs sur l'axe 2 et l'axe 3 respectivement.

Nous allons déterminer la pertinence de ces vecteurs formes en fonction des règles de décision, à savoir la règle des  $k$  –ppv et la discrimination par le calcul direct des frontières linéaires.

### 3.4.2 Décision par la règle des $k$ – ppv

Le choix de  $k$ , contrairement au raisonnement établi au paragraphe 2.4.2.1, est souvent arbitraire. En pratique, il est préférable de prendre  $k \ll N$ . Ainsi, nous choisirons  $k = 5$  pour la suite, bien qu'il soit possible de faire une étude sur  $k$ .

Reste à faire un choix pour le calcul de la distance dans cette règle de décision. La métrique utilisée dans le calcul de distance est très importante car, selon la forme des classes, des erreurs de classification peuvent être commises. Nous présentons, sur un exemple de données simulées, les deux cas de figures suivants :

- La distance euclidienne classique :

$$d_E(\mathbf{x}, \mathbf{y}) = [(\mathbf{x} - \mathbf{y})^t (\mathbf{x} - \mathbf{y})]^{1/2}$$

- La distance de Mahalanobis qui utilise l'inverse de la matrice de variance – covariance :

$$d_M(\mathbf{x}, \mathbf{y}) = [(\mathbf{x} - \mathbf{y})^t \mathbf{S}^{-1} (\mathbf{x} - \mathbf{y})]^{1/2}$$

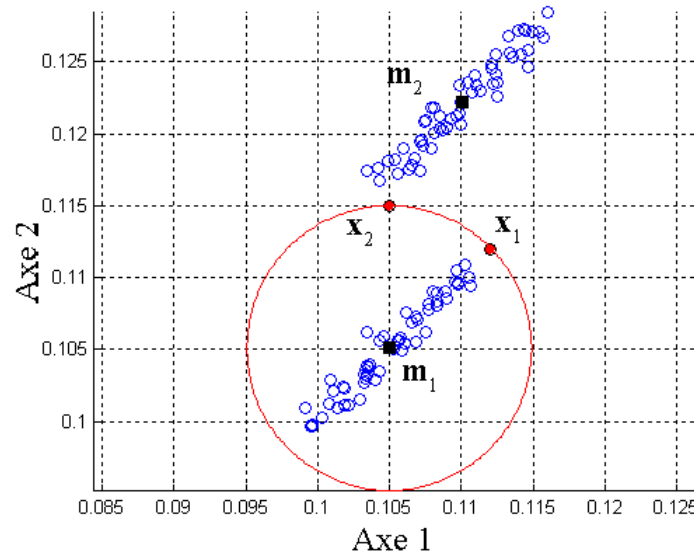
Le but est de classer deux nouveaux points par rapport à deux classes de dispersions quasiment identiques à celles de la plupart des classes montrées précédemment. Nous analyserons simplement les distances entre ces points et les barycentres des classes. L'illustration de la classification des deux points est faite sur la figure 3.26.

Le cercle indique que les deux points  $\mathbf{x}_1$  et  $\mathbf{x}_2$  sont à égale distance "Euclidienne" de la première

classe. Mais  $\mathbf{x}_2$  est encore plus proche de la seconde classe. Donc en toute logique, la décision est :

$$\begin{cases} \mathbf{x}_1 \rightarrow \omega_1 \\ \mathbf{x}_2 \rightarrow \omega_2 \end{cases}$$

Pourtant des erreurs aurait pu être commises, si l'on tient compte des résultats du tableau 3.7.



**Fig. 3. 26.** Illustration de la classification après calcul de la distance de Mahalanobis et de la distance Euclidienne.

	$\mathbf{m}_1$	$\mathbf{m}_2$	Décision
$d_M(\mathbf{x}_1, \mathbf{m}_i)$	0.0073	0.0219	Première classe
$d_M(\mathbf{x}_2, \mathbf{m}_i)$	0.0183	0.0074	Seconde classe
$d_E(\mathbf{x}_1, \mathbf{m}_i)$	0.0099	0.0104	<i>Première classe</i> <i>Risque d'erreur</i>
$d_E(\mathbf{x}_2, \mathbf{m}_i)$	0.0099	0.0088	Seconde classe

**Tableau. 3. 7.** Résultat des mesures de distance pour l'exemple de la figure 3.26

La distance Euclidienne entre le point  $\mathbf{x}_1$  et la seconde classe n'est pas nettement supérieure à sa distance Euclidienne par rapport à la première classe. Or la différence est parfaitement établie avec le calcul de la distance de Mahalanobis. Nous retiendrons cette dernière par la suite, car elle tient compte de la dispersion des classes.



Pour évaluer le système de décision basé sur la règle des  $k - \text{ppv}$  ( $k = 5$ ) et le vecteur forme de dimension 6, nous nous utilisons un ensemble de test tel que défini au paragraphe 3.4.

Cet ensemble de test est en quelque sorte variable car nous utilisons la méthode de "leave one out" ou plutôt du "leave 5 out". La procédure est la suivante :

Le vecteur forme est fixé. Pour chaque mode de fonctionnement 10 échantillons sont tirés aléatoirement et les 5 restants vont dans l'ensemble de test. A chaque itération (tirage aléatoire) l'espace de décision est redéfini en fonction des nouvelles données d'apprentissage. De cette manière, les ensembles d'apprentissage et de test sont toujours disjoints.

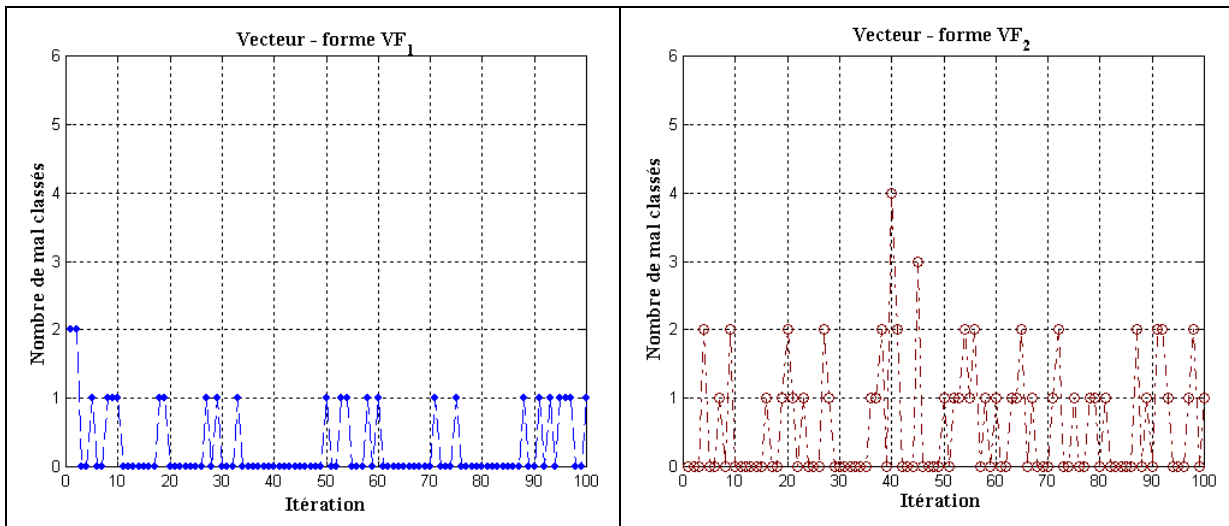
Après chaque sélection des éléments de l'ensemble d'apprentissage  $\mathbf{X}_a$  et de l'ensemble de test  $\mathbf{X}_t$  le nombre d'échantillons mal classés est calculé. Cela permet d'avoir une estimation des taux moyens de mauvaise classification :

$$T_M = \frac{\overline{N_M}}{N_t} \quad (3.33)$$

Avec :  $\overline{N_M}$  la moyenne du nombre de mal classés sur les cents itérations et  $N_t$  le nombre d'échantillons de test. La figure 3.27 montre le nombre d'observations mal classées à chaque itération.

Au final, le critère de défaut vaut respectivement 0,23% et 0,51% pour les ensembles d'apprentissages définis par les vecteurs formes  $VF_1$  et  $VF_2$ . Les très faibles pourcentages d'erreurs de classification prouvent donc la pertinence de ces deux signatures.

Cela signifie inversement que, sur la base de ces signatures, les performances diagnostic sont proche de 100%. Pour les deux cas de vecteur forme, les résultats d'affectation des observations de test sont repris en introduisant les rejets d'ambiguïté et de distance.

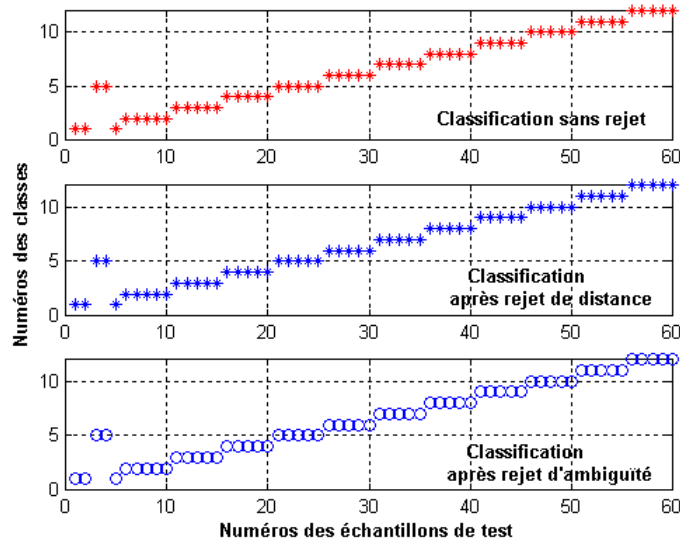


**Fig. 3.27.** Résultat de test sur les deux ensembles d'apprentissage. Les 60 échantillons de test sont sélectionnés aléatoirement.

Le rejet de distance est appliqué de manière locale. En effet, le seuil de rejet de distance est défini pour une classe  $\omega_i$  donnée en calculant les distances entre les points de celle – ci et son centre de gravité  $\mathbf{m}_i$ . Le seuil de rejet peut donc être assimilé au diamètre de la classe :

$$T_{di} = 2 \cdot \max_{j=1, N_i} [d(\mathbf{x}_{ij}, \mathbf{m}_i)] \tag{3.34}$$

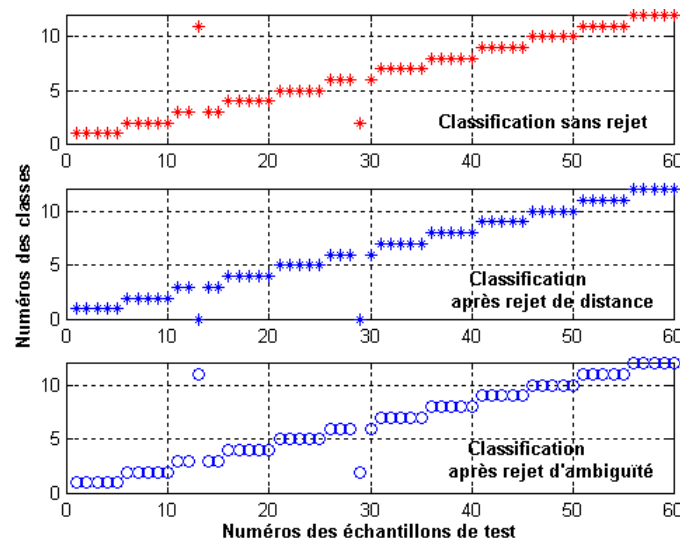
D'autre part, le nombre minimal de voisin  $k'$  que doit avoir un vecteur  $\mathbf{x}$  dans la classe la mieux représentée est fixé à  $k/2$ . Concrètement étant donné que nous avons choisi  $k = 5$ , le nombre minimal de voisin est 3. Les figures 3.28 et 3.29 présentent les résultats dans les trois cas, à savoir la classification sans rejet, après rejet de distance et après rejet d'ambiguïté. La mise à zéro signifie que le vecteur est rejeté.



**Fig. 3. 28.** Classification avant et après options de rejets des échantillons de test. Les vecteurs sont définis par  $VF_1$  (rapport de Fisher)

Sur la figure 3.28, Les deux erreurs de classification entre les classes  $\omega_1$  (machine saine à vide) et  $\omega_5$  (défaut 3 barres cassées à vide) montrent toute la difficulté de distinguer le défaut rotor pour de faibles niveaux de charge.

Sur la figure 3.29, les deux erreurs commises (observations n° 13 et n° 29) pour les observations définies par le vecteur forme  $VF_2$  peuvent vraiment être considérées comme étant des erreurs de diagnostic. Ce problème est corrigé par le rejet de distance.



**Fig. 3. 29.** Classification avant et après options de rejets des échantillons de test. Les vecteurs sont définis par  $VF_2$  (algorithme SBS)

La robustesse des signatures dans le cadre de la règle des  $k - ppv$  étant démontrée, il s'agira donc par la suite de prendre une décision pour toute nouvelle observation recueillie sur la machine.

**3.4.2.1 Phase d'exploitation : classification sans rejet**

Le diagnostic final à partir de la règle des  $k - ppv$  est établi pour les nouvelles observations  $X_u$  suivantes :

- Les mesures machine saine à 100 % de charge, notées  $C_1$
- Les mesures pour le défaut trois barres cassées à 100 % de charge, notées  $C_2$
- Les mesures pour une barre partiellement cassée de 0 % à 75 % de charge notées  $C_3$  à  $C_6$

Quinze enregistrements ont été effectués pour chacun de ces nouveaux modes de fonctionnement. Dans le cadre de la règle des  $k - ppv$  ( $k = 5$ ), nous allons appliquer la classification sans rejet à ces nouvelles observations.

Le tableau 3.8 donne les affectations des nouvelles observations aux classes de l'ensemble d'apprentissage définies par le vecteur forme  $VF_1$ .

Observations	$C_1$ Sain 100%	$C_2$ 3 barres 100%	$C_3$ 1 barre partielle 0%	$C_4$ 1 barre partielle 25%	$C_5$ 1 barre partielle 50%	$C_6$ 1 barre partielle 75%
Classification	15/15 $\rightarrow \omega_8$	15/15 $\rightarrow \omega_8$	3/15 $\rightarrow \omega_1$ 9/15 $\rightarrow \omega_5$ 3/15 $\rightarrow \omega_9$	7/15 $\rightarrow \omega_2$ 7/15 $\rightarrow \omega_5$ 1/15 $\rightarrow \omega_6$	10/15 $\rightarrow \omega_3$ 3/15 $\rightarrow \omega_5$ 2/15 $\rightarrow \omega_{11}$	15/15 $\rightarrow \omega_4$

**Tableau. 3. 8.** Répartition des nouvelles observations dans l'espace défini par  $VF_1$

La classification des observations pour une barre partiellement cassée aux classes "machine saine" ou "défaut trois barres cassées" serait admissible. En effet, il s'agit d'un défaut de faible niveau de gravité.

L'affectation des nouvelles observations représentées par  $C_1$  à la classe  $\omega_8$  (défaut 3 barres cassées à 75 % de charge) est un échec du diagnostic. L'affectation d'une partie des points de  $C_4$  et  $C_5$  à la classe  $\omega_5$  (défaut 3 barres cassées à 0% de charge) est également surprenante. En effet, au vu de la disposition des classes de l'ensemble d'apprentissage (figure 3.24), l'espace de

représentation est fortement conditionné par le niveau de charge. Or les mesures  $C_4$  et  $C_5$  représentent le défaut rotor "une barre", à 25 % et 50 % de charge respectivement.

Enfin la répartition des points de  $C_3$  sur les trois classes à vide montre toute la difficulté de distinguer ces états de la machine pour des faibles niveaux de charge.

Sur l'ensemble défini par le vecteur forme  $VF_2$ , les résultats de classification sont nettement plus acceptables (tableau 3.9)

Observations	$C_1$ Sain 100%	$C_2$ 3 barres 100%	$C_3$ 1 barre partielle 0%	$C_4$ 1 barre partielle 25%	$C_5$ 1 barre partielle 50%	$C_6$ 1 barre partielle 75%
Classification	15/15 $\rightarrow \omega_4$	15/15 $\rightarrow \omega_8$	2/15 $\rightarrow \omega_5$ 13/15 $\rightarrow \omega_9$	15/15 $\rightarrow \omega_2$	1/15 $\rightarrow \omega_3$ 14/15 $\rightarrow \omega_7$	15/15 $\rightarrow \omega_4$

**Tableau. 3. 9.** Répartition des nouvelles observations dans l'espace défini par  $VF_2$

Les mesures "machine saine" et "défaut 3 barres cassées"  $C_1$  et  $C_2$  sont bien identifiées.

Les points de  $C_3$  sont classés en défaut rotor et en défaut stator à vide. Cela confirme la difficulté d'identifier des nouveaux modes pour de faibles niveaux de charge. Par ailleurs, les points de  $C_4$  et  $C_6$  sont identifiés en tant que données "machine saine" en accord avec le niveau de charge alors que pour l'ensemble  $C_5$ , la plupart des données sont affectées à  $\omega_7$ , c'est à dire la classe représentant le défaut trois barres cassées à 50 % de charge.

Le fonctionnement avec une barre partiellement cassée est proche du fonctionnement sain même s'il s'agit d'un début de mode de défaillance. Cela explique la confusion fréquemment faite entre ces deux modes.

A l'issue de cette phase de diagnostic sans rejet, un premier constat peut être fait : le vecteur forme défini par l'algorithme SBS donne un résultat beaucoup plus logique en ce qui concerne l'identification des nouvelles observations.

Néanmoins, la classification sans rejet ne permet pas une détection de nouveaux modes de fonctionnement possibles.

Les options de rejet permettrons, en quelque sorte, de quantifier l'appartenance des nouvelles observations à leurs classes d'affectation.

### 3.4.2.2 Phase d'exploitation : classification avec rejet

Nous rappelons que le rejet d'ambiguïté reflète l'indécision éventuelle entre plusieurs classes et le rejet de distance traduit le fait que la nouvelle observation du système ne correspond à aucune des classes connues. Notons  $\omega_0$  la classe des observations rejetées en ambiguïté et  $\omega_d$  celle des observations rejetées en distance.

Les tableaux suivants reprennent les résultats de classification après application des rejets de distance et d'ambiguïté. La décision, en tenant compte des deux options de rejet, est exprimée par l'équation (2.64) dans le chapitre 2. Lorsqu'une observation est rejetée à la fois en distance et en ambiguïté, il est logique de ne prendre en compte que le rejet de distance. Les résultats sont les suivants :

Observations	<b>C<sub>1</sub></b> Sain (100%)	<b>C<sub>2</sub></b> 3 barres (100%)	<b>C<sub>3</sub></b> 1 barre partielle (0%)	<b>C<sub>4</sub></b> 1 barre partielle (25%)	<b>C<sub>5</sub></b> 1 barre partielle (50%)	<b>C<sub>6</sub></b> 1 barre partielle (75%)
Rejet de distance	15/15	15/15	0/15	14/15	15/15	15/15
Rejet d'ambiguïté	0/15	0/15	0/15	9/15	6/15	0/15
Décision	15/15 → $\omega_d$	15/15 → $\omega_d$	3/15 → $\omega_1$ 9/15 → $\omega_5$ 3/15 → $\omega_9$	14/15 → $\omega_d$ 1/15 → $\omega_0$	15/15 → $\omega_d$	15/15 → $\omega_d$

**Tableau. 3. 10.** *Décision avec options de rejet pour les nouvelles observations dans l'espace défini par  $VF_1$*

Observations	C <sub>1</sub> Sain (100%)	C <sub>2</sub> 3 barres (100%)	C <sub>3</sub> 1 barre partielle (0%)	C <sub>4</sub> 1 barre partielle (25%)	C <sub>5</sub> 1 barre partielle (50%)	C <sub>6</sub> 1 barre partielle (75%)
Rejet de distance	15/15	15/15	2/15	15/15	15/15	15/15
Rejet d'ambiguïté	0/15	0/15	0/15	0/15	10/15	0/15
Décision	15/15 → ω <sub>d</sub>	15/15 → ω <sub>d</sub>	13/15 → ω <sub>g</sub> 2/15 → ω <sub>d</sub>	15/15 → ω <sub>d</sub>	15/15 → ω <sub>d</sub>	15/15 → ω <sub>d</sub>

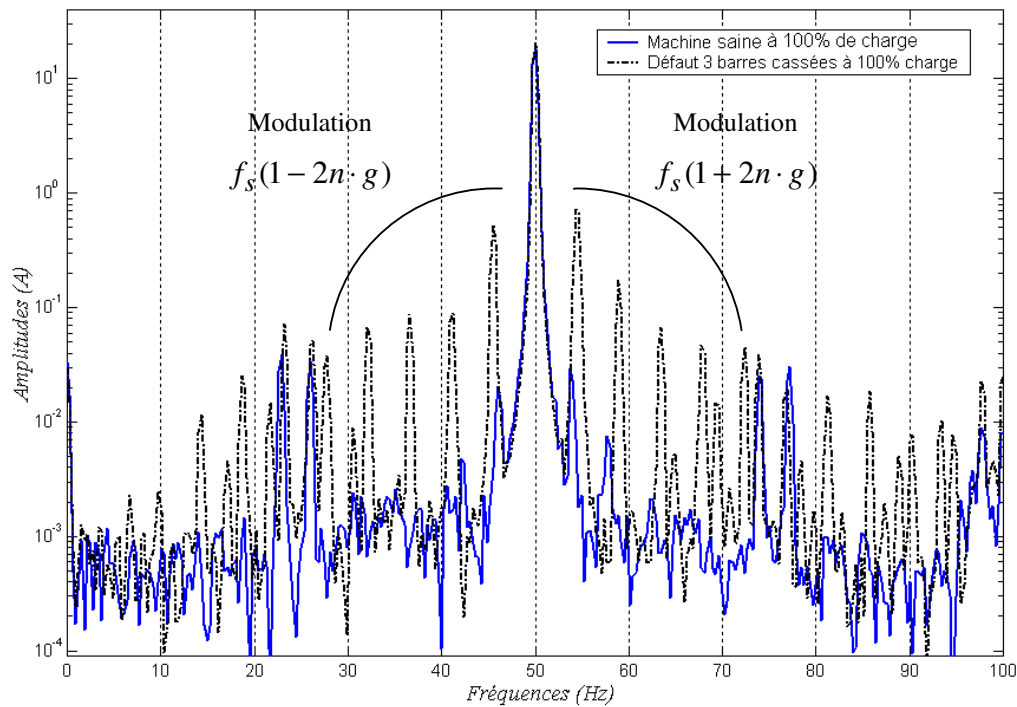
**Tableau. 3. 11.** *Décision avec options de rejet pour les nouvelles observations dans l'espace défini par VF<sub>2</sub>*

Pour les deux espaces de représentation, le rejet de distance n'est pratiquement jamais appliqué aux observations à vide.

Le système a soit différé la décision (ambiguïté), soit rejeté des observations à cause de leurs faibles correspondances avec les classes connues. Dans ce dernier cas, le diagnostic de la machine est insuffisant.

Reste donc à analyser ces signaux à l'aide des procédures classiques de traitement. Les signatures des défauts de la machine asynchrone étant pour la plupart connues (chapitre 1), l'identification de ces nouvelles mesures est donc possible.

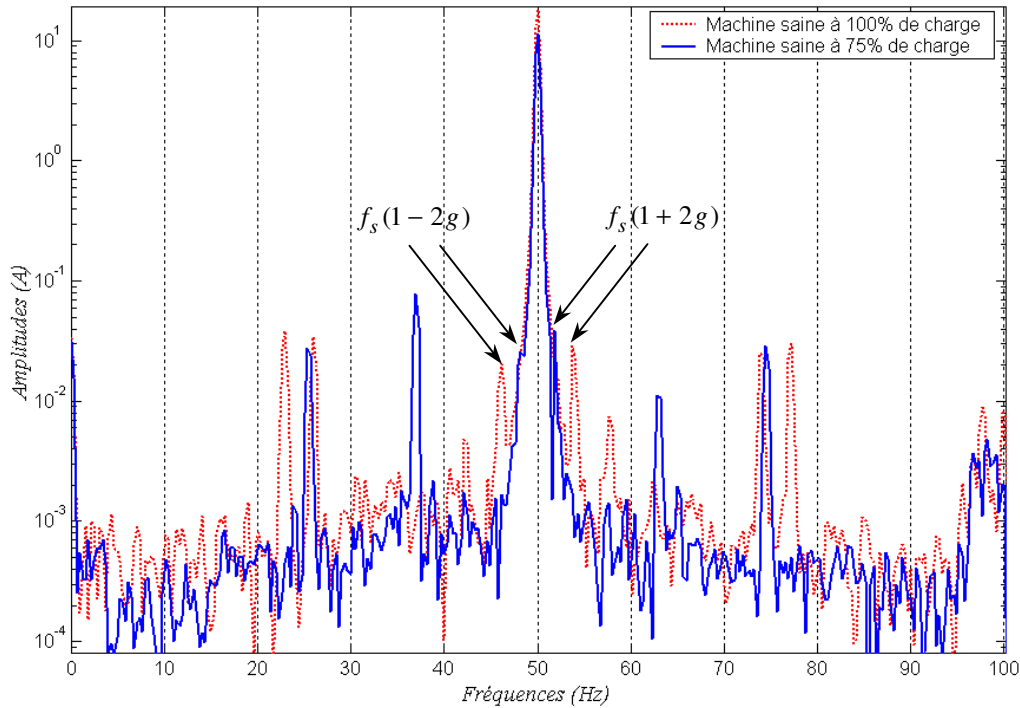
Supposons que l'utilisateur ait effectué ces mesures à des instants bien précis en connaissant les conditions de fonctionnement, à savoir la vitesse de rotation, la puissance utile en bout d'arbre ou encore les valeurs des courants de ligne. Il suffit de faire une analyse fréquentielle, temporelle ou paramétrique (estimation des résistances, inductance etc...) pour savoir si ces mesures correspondent à des modes de défaillances ou non. Si l'on prend l'exemple des mesures "machine saine à pleine charge" et "défaut 3 barres cassées à pleine charge", l'identification du défaut est indiscutable car les raies au fréquences  $f_s(1 \pm 2n \cdot g)$  traduisent une forte modulation du courant de ligne (figure 3.30).



**Fig. 3.30.** Analyse spectrale des mesures "Machine saine" et "3 barres cassées" à 100 % de charge

Or la comparaison entre les mesures "machine saine" à 75 % et 100 % de charge montre que la différence majeure entre les spectres (figure 3.31) tient du niveau de charge. En effet, le glissement étant plus important à pleine charge, les raies à  $f_s(1 \pm 2g)$  sont un peu plus éloignées du fondamental.





**Fig. 3.31.** Analyse spectrale des mesures "Machine saine" à 75 % et à 100 % de charge

Les points rejetés pourront ensuite être insérés dans l'ensemble d'apprentissage. Il est possible également de lancer un algorithme de coalescence pour rechercher une structure en classes des points rejetés. Mais la coalescence n'est pas obligatoire, si les modes correspondants sont connus et bien séparés.

La procédure de diagnostic pourra donc être établie selon le schéma de principe de la figure 3.32.

Lors de l'étape d'actualisation il n'est pas nécessaire de refaire une sélection de paramètres pour mieux définir l'ensemble d'apprentissage augmenté. En effet il n'est pas obligatoire de recalculer un vecteur forme qui aurait bien caractérisé les points ainsi que les classes auparavant. Par contre, il faudra redéfinir les matrices de variance intra – classe et inter – classes.

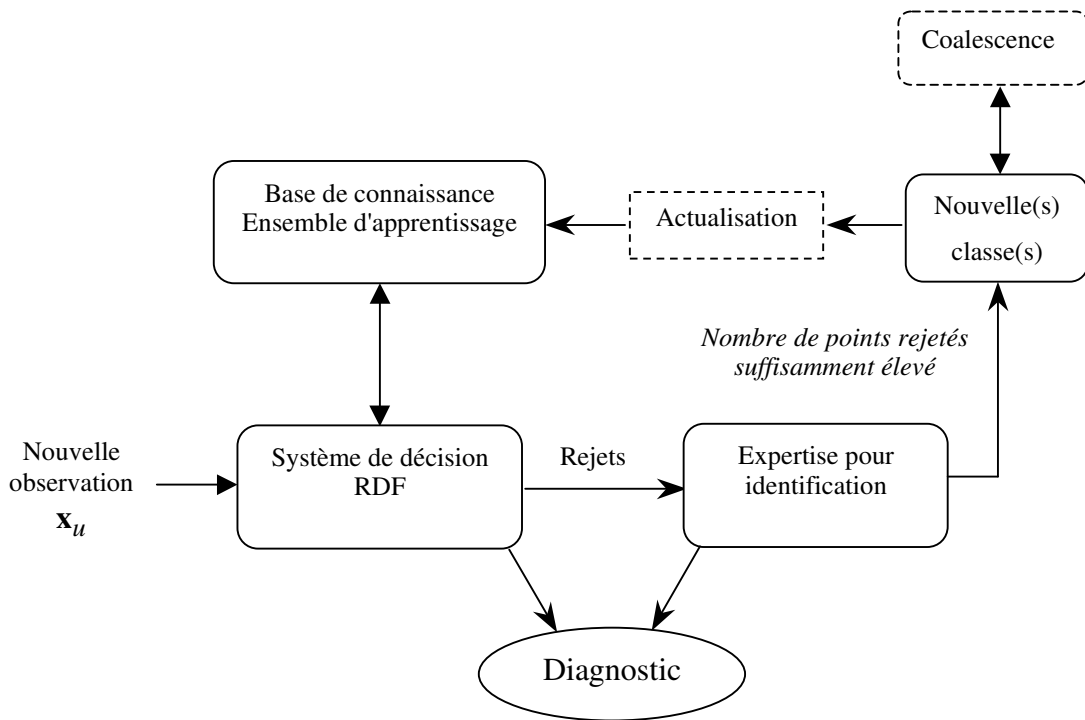


Fig. 3. 32. Schéma de principe de la procédure de diagnostic

Les résultats de classification montrent que le vecteur forme donnant les résultats les plus cohérents est  $VF_2$ . Cependant, les erreurs commises en assimilant les mesures "1 barre partielle à vide" à  $\omega_0$  (défaut de déséquilibre stator à vide) n'ont pas été corrigées par les options de rejet.

La règle de décision, basée sur le calcul direct de frontières linéaires entre les classes a également été développée et utilisée dans la procédure de décision.

### 3.4.3 Décision à partir des frontières linéaires

La règle des  $k$  – plus proches voisins repose sur la détermination de frontières décrites par des propriétés statistiques. Elle ne tient pas compte, en réalité, de la position des classes et des échantillons mais des distances entre les observations et les classes de l'ensemble d'apprentissage.

Le calcul des frontières linéaires et non linéaires a pour objectif de délimiter les zones de l'espace correspondantes aux classes. Ainsi, une nouvelle observation sera affectée à une classe selon sa position par rapport à cette dernière et non directement par rapport à sa distance.

D'autre part, comme nous l'avons montré au cours du chapitre 2, la prise de décision est beaucoup plus rapide dans la mesure où seuls les coefficients des équations mathématiques des frontières sont conservés. Nous verrons toutefois que d'autres paramètres de décision vont être pris en compte pour

introduire le rejet de distance. Ces paramètres sont les seuils propres à chaque classe  $T_{di}$  définis par l'équation (3.34) ainsi que les centres de gravité des classes.

Le séparateur linéaire utilisé est le discriminateur de Fisher.

Nous rappelons que les vecteurs poids  $\mathbf{w}$  et les valeurs de seuil  $w_0$  sont définis pour chaque dichotomie à l'aide des éléments constituant les classes.

Ainsi pour une décision entre deux classes  $\omega_i$  et  $\omega_j$  la frontière sera défini par :

$$\mathbf{w}^{(ij)} = \left( \mathbf{S}_w^{(ij)} \right)^{-1} (\mathbf{m}_i - \mathbf{m}_j) \quad (3.35)$$

Et

$$w_0^{(ij)} = -\mathbf{m}_{ij}^t \mathbf{w}^{(ij)} \quad (3.36)$$

$\mathbf{S}_w^{(ij)}$  et  $\mathbf{m}_{ij}$  sont respectivement la matrice de variance intra – classe et le centre de gravité de l'ensemble des deux classes. La décision entre  $M$  classes, schématisée par l'arbre de décision de la figure 2.17, peut être reprise de manière algorithmique :

---

*Initialisation* :  $i = 1, j = 2$

**For**  $j = 2, M$

$$g(\mathbf{x}) = (\mathbf{w}^{(ij)})^t \mathbf{x} + w_0^{(ij)}$$

**If**  $g(\mathbf{x}) > 0$

$\mathbf{x} \in \omega_i$

**Else**

$\mathbf{x} \in \omega_j$

$i = j$

**End**

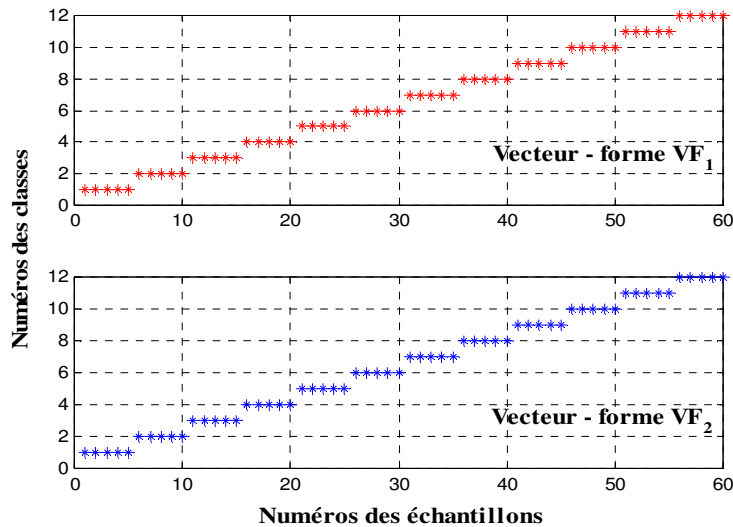
**End**

---

### 3.4.3.1 Résultats de classification

La figure 3.33 montre les résultats de classification pour les individus de test. Toutes les observations sont correctement identifiées. Ce résultat montre que les classes sont linéairement séparables.

Nous ne retrouvons pas les erreurs commises lors de l'application de la règle des  $k - ppv$  pour les modes de fonctionnement à vide.

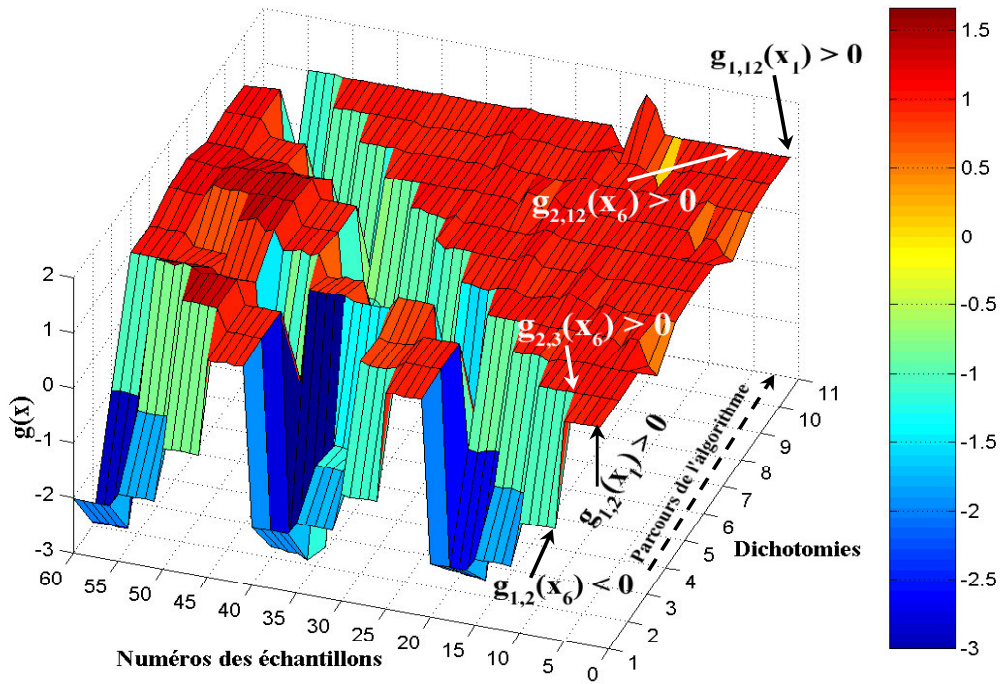


**Fig. 3. 33.** Résultats du diagnostic sans rejet, par les frontières linéaires, effectué sur les observations de test.

La figure 3.34 illustre le parcours complet de l'algorithme pour chacun des échantillons. Les dichotomies  $i$  correspondent aux décisions prises entre la classe  $\omega_i$  et la classe  $\omega_j$ . A titre d'exemple, si nous considérons l'observation  $\mathbf{x}_6$  alors :

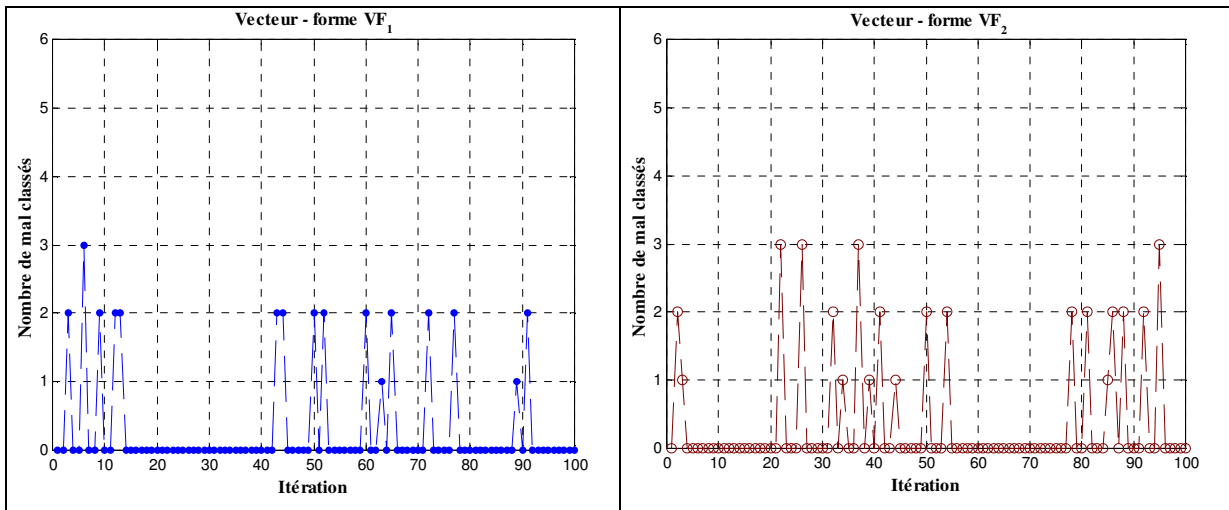
Après la première dichotomie entre  $\omega_1$  et  $\omega_2$ , on obtient  $g_{1,2}(\mathbf{x}_6) < 0$  donc  $\mathbf{x}_6 \notin \omega_1$ . Mais le test doit être pratiqué entre  $\omega_2$  et les autres classes. Au final,  $g_{2,12}(\mathbf{x}_6)$  étant positif, l'observation  $\mathbf{x}_6$  est affectée à la classe  $\omega_2$ .

Ainsi  $(M - 1)$  valeurs sont calculées pour chaque observation, alors que pour la règle des  $k - ppv$ , il faut calculer les distances entre ces vecteurs et tous les vecteurs de l'ensemble d'apprentissage. Cette règle permet un gain important en temps de calcul.



**Fig. 3. 34.** Récapitulatif du calcul effectué pour chaque observation vis à vis de sa position par rapport à l'ensemble des classes.

Mais si le diagnostic de ces individus de test est parfait, il convient comme précédemment, de faire plusieurs fois une sélection aléatoire de cinq échantillons parmi quinze pour chaque classe. Cela permet de juger la robustesse du tracé de frontières entre les classes et la pertinence des signatures. Le résultat de cette sélection est présenté sur la figure 3.35.

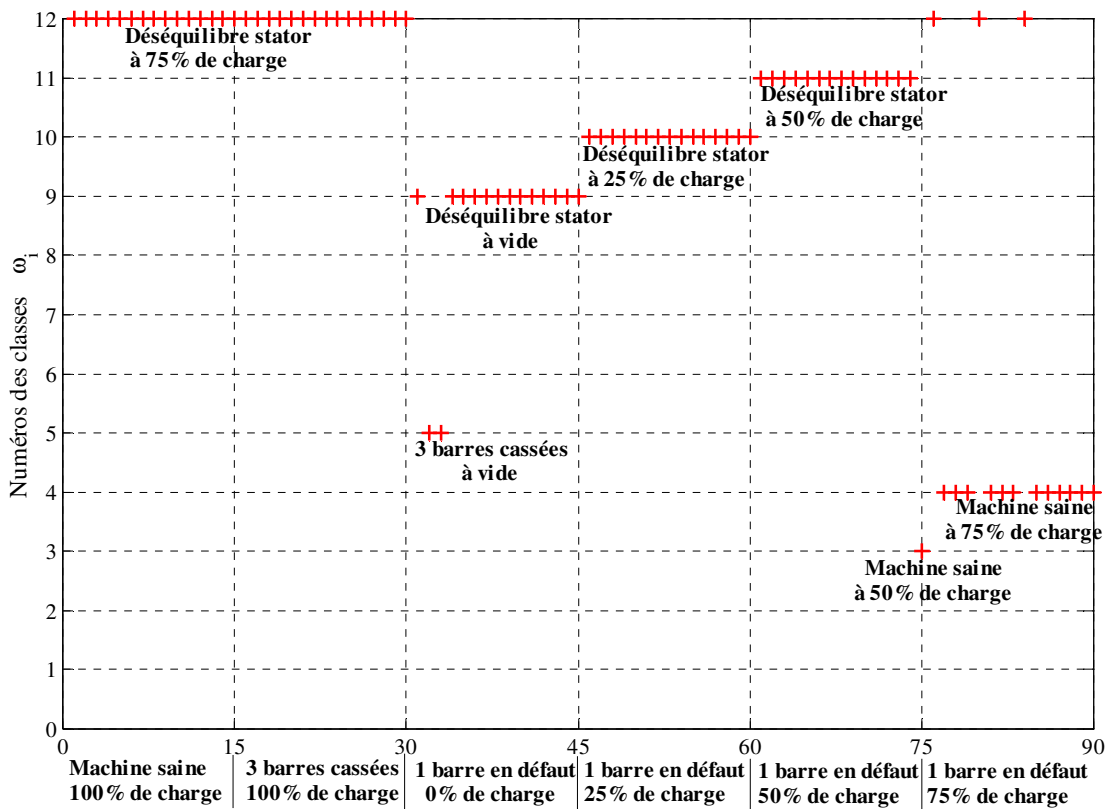


**Fig. 3. 35.** Résultats en terme de nombre d'échantillons mal classés après cent sélections aléatoires des individus de test. Décision par calcul direct des frontières.

Les pourcentages d'erreur sont respectivement de 0,27 % et de 0,316% pour  $VF_1$  et  $VF_2$ . La discrimination linéaire offre donc les mêmes performances de diagnostic (en tout cas pour des modes connus) que la règle des  $k - ppv$ .

Pour l'analyse des nouvelles observations nous avons choisi de tester cette règle de décision à partir de  $VF_2$ . En effet, pour ce vecteur forme, les résultats de diagnostic avant rejet permettaient une meilleure identification des modes correspondants aux nouvelles mesures.

Les décisions sans rejet prises pour les nouvelles mesures sont répertoriées sur la figure suivante. Les classes auxquelles sont affectés les points sont précisées sur la figure.



**Fig. 3. 36.** Phase de décision sans rejet par les frontières linéaires pour les nouvelles observations. L'espace est caractérisée par  $VF_2$

Ces données montrent que, s'il est possible d'obtenir les vecteurs – poids définissant les hyperplans, nous n'avons par contre aucune garantie sur la fiabilité du diagnostic des nouvelles mesures. En effet les observations sont classées selon leurs positions dans l'espace et beaucoup moins selon leurs distances aux classes.

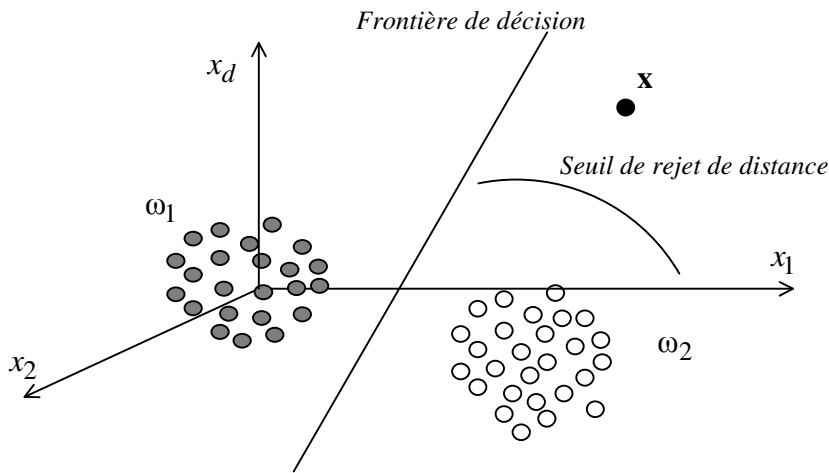
L'utilisateur ne peut donc pas considérer ces seules informations pour diagnostiquer l'état de la

machine en fonction des nouvelles mesures. D'où la nécessité de faire bénéficier à l'algorithme de décision des options de rejets pour différer la classification ou pour interpréter l'appartenance des observations aux modes connus.

### 3.4.3.2 Extension à la notion de rejet.

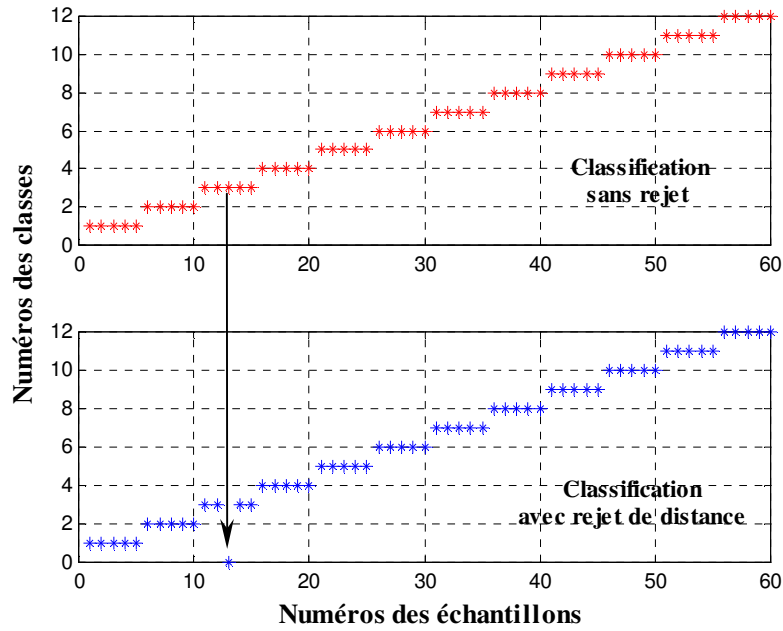
Les seuils de distance relatifs aux diamètres des classes sont utilisés également dans cette règle de décision.

Une frontière linéaire séparant deux classes divise l'espace en deux zones égales. Une observation correspondant à une classe donnée sera logiquement affectée à celle – ci. Mais le point doit être jugé suffisamment proche de la classe pour y être réellement affecté (figure 3.37).



**Fig. 3. 37.** Illustration de la discrimination linéaire avec rejet de distance

L'application du rejet de distance aux observations de test caractérisées par le vecteur forme  $VF_2$  est présentée sur la figure 3.38. Pour l'observation n° 13 (machine saine à 50 % de charge), le résultat est identique à celui obtenu par les  $k - ppv$ . Le point est rejeté en distance car il est trop éloigné par rapport à la classe  $\omega_3$ . Par contre, avec la classification sans rejet, l'erreur précédente qui a été commise en assimilant cette observation au mode (déséquilibre stator à 50% de charge) représenté par la classe  $\omega_{11}$  n'est pas réitérée. Le diagnostic est donc sensiblement meilleur.



**Fig. 3. 38.** Résultats de classification par les frontières linéaires après application du rejet de distance. Les observations de test sont définies par  $VF_2$ .

Par contre, pour les nouvelles observations, la décision après application du rejet de distance, est quasiment identique aux résultats des  $k - ppv$  si l'on se réfère au tableau suivant :

Observations	C <sub>1</sub> Sain 100%	C <sub>2</sub> 3 barres 100%	C <sub>3</sub> 1 barre partielle 0%	C <sub>4</sub> 1 barre partielle 25%	C <sub>5</sub> 1 barre partielle 50%	C <sub>6</sub> 1 barre partielle 75%
Rejet de distance	15/15	15/15	2/15	15/15	15/15	15/15
Décision	15/15 → $\omega_d$	15/15 → $\omega_d$	13/15 → $\omega_g$ 2/15 → $\omega_d$	15/15 → $\omega_d$	15/15 → $\omega_d$	15/15 → $\omega_d$

**Tableau. 3. 12.** Résultats de la classification des nouvelles observations par les frontières linéaires avec rejet de distance. Les observations sont définies par  $VF_2$ .

Toutefois, ce ne sont pas exactement les mêmes observations de l'ensemble C<sub>3</sub> qui sont rejetées en distance.

Nous établirons une comparaison de ces deux résultats au cours du paragraphe suivant.

Le rejet d'ambiguïté n'a pas été traité dans le cadre de la discrimination linéaire. Toutefois celui –



ci est possible.

Comme ce fut précisé dans le paragraphe 2.4.3.3, le rejet d'ambiguïté peut être appliqué en définissant une marge de sécurité  $\mathcal{E}$  telle que :  $\mathbf{x} \rightarrow \omega_0$  si  $|g(\mathbf{x})| \leq \mathcal{E}$ . Mais il est également possible d'avoir recours aux distances entre les observations et les frontières. Ces distances sont connues (figure 2.14) :

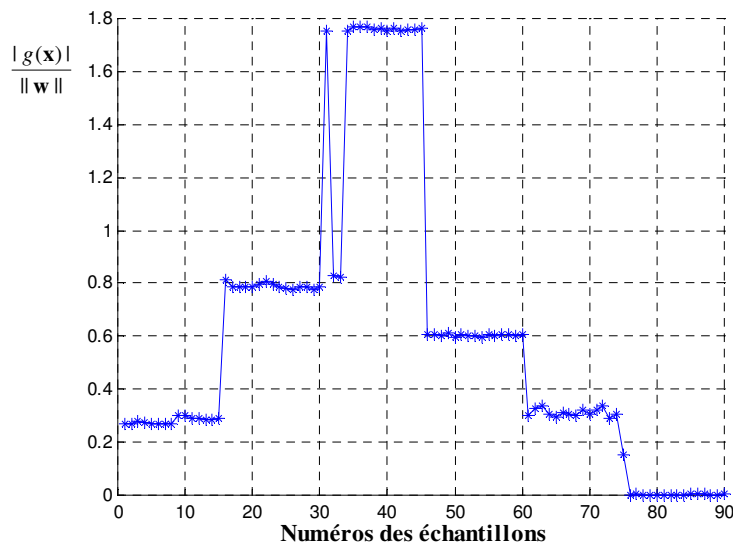
$$d_{\mathbf{x},H} = \frac{|g(\mathbf{x})|}{\|\mathbf{w}\|} \quad (3.37)$$

Où  $\|\mathbf{w}\|$  désigne la norme de  $\mathbf{w}$ .

C'est l'approche utilisée par [Didelet, 92] en supposant les classes gaussiennes. L'auteur démontre, dans le cadre de ses travaux basés sur les arbres de neurones avec rejet d'ambiguïté, que si deux classes ont leurs matrices de variance – covariance identiques, alors la valeur optimale du seuil de rejet appliqué à  $d_{\mathbf{x},H}$  est :

$$T_a = \frac{\ln(2)}{\|\mathbf{w}\|} \quad (3.38)$$

Le tracé (figure 3.39) des distances entre les nouvelles observations et les hyperplans ayant servi à les classer, devrait donner des résultats intéressants pour des travaux futurs, car certaines observations sont très proches des frontières.



**Fig. 3. 39.** Calcul des distances entre les nouvelles observations et les hyperplans

### 3.4.4 Comparaison des procédures de décision

Les résultats précédents ont mis en lumière les performances de diagnostic des deux algorithmes de décision. Pour les modes de fonctionnement connus, les deux algorithmes de décision offrent des résultats fiables pratiquement à cent pour cent. Nous retiendrons la facilité d'élaboration de ces règles de décision et surtout le gain de temps offert par le calcul direct des frontières.

Par contre pour les nouvelles mesures les résultats sont sensiblement différents. Les options de rejets, en particulier le rejet de distance, ont aidé à établir la correspondance ou la non-appartenance de ces nouvelles observations aux modes de fonctionnement connus.

Parmi ces nouvelles observations, certaines représentaient les états de la machine déjà appris mais pour des niveaux de charge différents. Il s'agit des mesures "machine saine" et "machine avec trois barres cassées au rotor" à pleine charge. Par conséquent il serait logique, qu'avant rejet, les algorithmes de décision identifient correctement ces modes.

Pourtant, seule la règle des  $k - ppv$  appliquée aux observations définies par le vecteur forme  $VF_2$  (vecteur des paramètres  $m_p, m_q, \sigma_\alpha, \sigma_\beta, \sigma_s$  et  $Z_d$  sélectionnés avec l'algorithme SBS) a donné satisfaction.

Le choix de la signature de la machine étant définitivement fixé, il est intéressant de voir les résultats des règles de décision pour des vecteurs définis par ces paramètres mais dans une situation différente : la machine alimentée par le biais d'un onduleur.

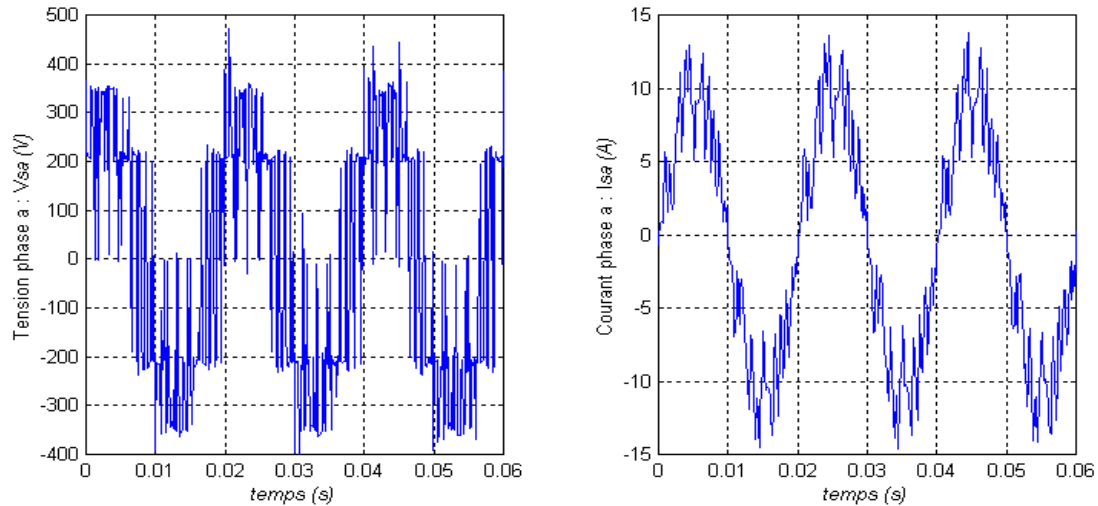
### 3.4.5 Cas de l'alimentation avec onduleur

Lorsque la machine asynchrone est alimentée par l'intermédiaire d'un convertisseur, les courants sont affectés par les harmoniques multiples de la fréquence de commutation. Les formes d'ondes des signaux (figure 3.40) ainsi que leur contenu fréquentiel rendent plus difficile la détection des modes défailants. L'étude de ce cas particulier a pour but de justifier la possibilité d'appliquer des méthodes de reconnaissance de formes au diagnostic de la machine avec ce type d'alimentation.

Nous rappelons que les différentes classes obtenues représentent à vide et en charge :

- la machine saine
- le cas du rotor ayant 4 barres cassées
- le déséquilibre d'alimentation du stator

Nous disposons de 15 acquisitions par mode de fonctionnement, en charge et à vide. Cet ensemble de données est bien distinct de l'ensemble d'apprentissage utilisé précédemment sans onduleur.

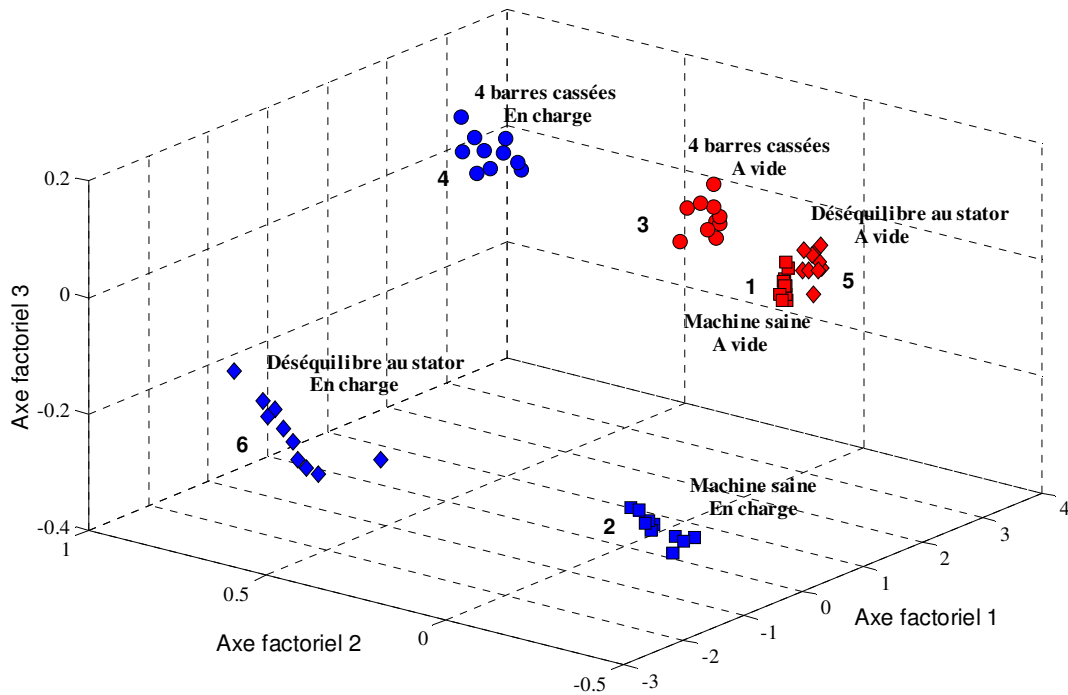


**Fig. 3.40.** Tension générée par l'onduleur sur une phase de la machine, et courant de ligne correspondant.

L'application de la règle des  $k - ppv$  a montré l'efficacité du vecteur forme  $VF_2$ . Nous conserverons ce même vecteur forme quelque soit l'ensemble d'apprentissage, tant que les taux d'erreurs resteront acceptables.

Nous allons donc tester ces règles de décisions pour les données onduleur avec ce vecteur forme  $VF_2$ .

La projection de ces données dans le sous – espace défini par l'ACP est présentée sur la figure 3.41. L'ensemble d'apprentissage est constitué de dix éléments par classe.

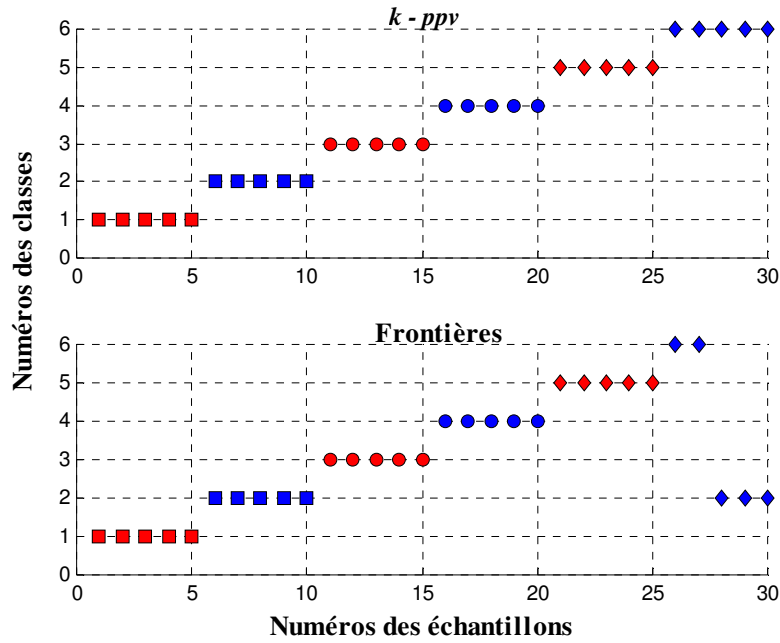


**Fig. 3. 41.** Visualisation des données onduteur, caractérisées par les paramètres de  $VF_2$ , dans le meilleur sous – espace de dimension 3. 99,8 % de l'information est conservée.

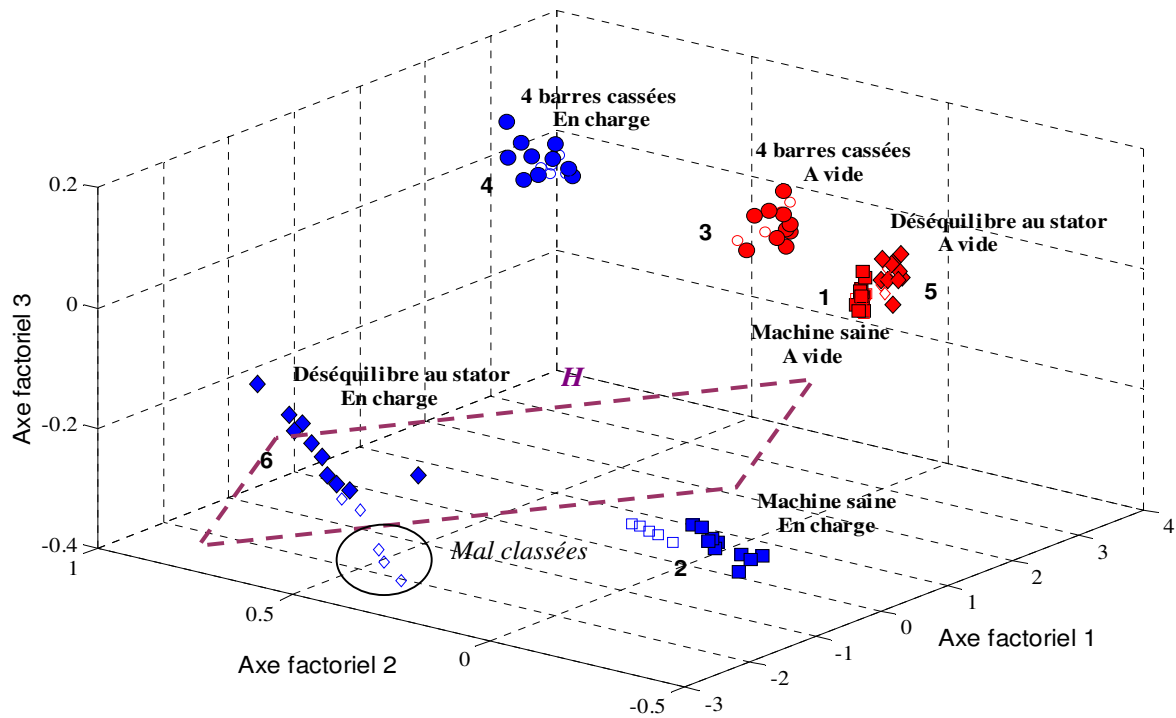
La classification est effectuée pour les cinq observations supplémentaires de chacune des classes, avec les deux règles de décision.

Les résultats de la figure 3.42 montrent que dans le cas de la discrimination linéaire, des erreurs de diagnostic sont commises entre la classe  $\omega_6$  (déséquilibre stator en charge) et la classe  $\omega_2$  (machine saine en charge). Or dans le cas de la règle des  $k - ppv$ , le diagnostic est correct.

Ces erreurs obtenues avec la discrimination linéaire peuvent s'expliquer si l'on suppose que la frontière séparant la classe  $\omega_2$  de la classe  $\omega_6$  est voisine du plan  $H$  (figure 3.43). La fonction  $g_{2,6}$  appliquée à ces échantillons étant positive, on comprend pourquoi ils sont mal classés.



**Fig. 3. 42.** Résultats de classification, par les  $k - ppv$  et le calcul des frontières, des individus supplémentaires issus des données onduleur.



**Fig. 3. 43.** Visualisation de l'ensemble des données onduleur dans l'espace factoriel. Le plan H donne la position éventuelle de la frontière séparant les classes 2 et 6.

Ces résultats montrent que ce vecteur forme peut aussi être conservé pour le diagnostic de la machine en alimentation onduleur.

---

### 3.4.6 Conclusion

Les résultats précédents ont mis en évidence les performances des algorithmes de décisions pour diagnostiquer les différents états de la machine. Ces bonnes performances, surtout lorsque les nouvelles observations correspondent aux modes déjà connus, ne sont obtenues que si la signature (autrement dit le vecteur forme) est correctement définie.

Nous avons décrit dans ce chapitre les paramètres permettant de caractériser les modes de fonctionnement. A partir de ces données, nous avons extrait les informations les plus pertinentes et construit le vecteur forme.

L'algorithme SBS ("Sequential Backward Selection") a donné, dans cette phase de construction, les meilleurs résultats. En effet, le diagnostic des observations définies par le vecteur forme construit avec cet algorithme était meilleur.

Nous avons testé les algorithmes de décision avec les rejets d'ambiguïté et de distance. Ces derniers permettent soit de différer la décision, soit de rejeter le point considéré.

Cette partie (possibilité de rejet) du système de diagnostic par reconnaissance des formes est très importante dans la mesure où elle permet d'enrichir la base de connaissance et de construire de nouvelles classes. Cela est possible, bien entendu, après l'analyse et l'identification complète des observations rejetées par l'utilisateur.

Le système de diagnostic devient alors évolutif.



---

## Conclusion générale

---

Les machines électriques synchrones, asynchrones et à courant continu nécessitent une détection rapide et précoce de leurs défaillances, à cause de leurs utilisations fréquentes dans diverses applications. Le travail présenté dans ce mémoire expose l'apport des méthodes de reconnaissance des formes au diagnostic des défauts des machines asynchrones.

Pour aborder l'étude, nous avons établi un état de l'art des défauts pouvant intervenir dans la machine asynchrone et des méthodes classiques permettant de les identifier. Ce travail préliminaire a mis en évidence deux catégories d'approches en vue du diagnostic spécifique de la machine asynchrone :

- le premier type d'approches consiste à utiliser des modèles analytiques du processus machine asynchrone. Ceux-ci fournissent une estimation en ligne des grandeurs non mesurables, inhérentes à la machine et variant en fonction des états de cette dernière.
- la seconde catégorie d'approches consiste à analyser les signaux mesurés sur la machine en cours de fonctionnement. Les analyses sont réalisées, pour la plupart, dans le domaine fréquentiel et en régime permanent. Les informations issues de ces analyses permettent souvent une identification efficace des défaillances présentes au stator, au rotor et sur les paliers (roulements). Toutefois ces approches fréquentielles trouvent leurs limites en présence d'importants bruits de mesure qui noient l'information pertinente.

Cette première étude est typiquement du ressort de l'expert du processus qu'il faut surveiller. Elle permet d'acquérir une connaissance substantielle de la machine asynchrone, de partir sur une première base de mesures directes ou indirectes accessibles sur la machine. Mais cette expertise ne résout pas en elle-même la problématique du diagnostic automatique des défauts ; elle sert de point de départ à la mise en place d'un éventuel système de décision.

La deuxième partie du travail de la thèse consiste à introduire la procédure de mise en œuvre d'un système de décision basé sur la reconnaissance des formes. Cette méthodologie



est générique. En effet, elle peut s'appliquer à tout système complexe (être humain, réseau téléphonique, véhicule automobile...) et n'est donc pas spécifique au problème de la machine asynchrone. Le mécanisme de mise en place d'un système de diagnostic à base de reconnaissance des formes nécessite l'existence d'une bibliothèque de mesures réalisées sur le processus et totalement étiquetées. Un ensemble d'apprentissage et un ensemble de test, totalement disjoints, en sont extraits. L'ensemble d'apprentissage est utilisé pour effectuer le choix des paramètres pertinents par rapport aux défauts recherchés. L'analyse en composantes principales, le critère de Fisher ou les méthodes de type "Sequential Backward Selection" donnent des indicateurs permettant de juger de la qualité des paramètres retenus. Ceux-ci sont rassemblés dans un vecteur forme, appelé aussi signature caractéristique du processus à surveiller.

Sur la base de cette signature sont alors construits un espace de représentation des échantillons mesurés sur le processus, puis un espace de décision permettant d'identifier une structure de classes caractérisant géométriquement les différents modes de fonctionnement de l'ensemble d'apprentissage. L'ajout d'une règle de décision permet de classer automatiquement une mesure ou observation, c'est à dire de décider automatiquement du mode de fonctionnement que représente cette mesure. L'algorithme des  $k$  – plus proches voisins ( $k$  – PPV) consiste à associer une mesure à la classe majoritairement représentée parmi les mesures voisines les plus proches (au sens d'une distance qui peut être euclidienne ou de Mahalanobis par exemple). Coûteux en temps de calcul, cet algorithme peut être remplacé par une approche consistant à partitionner l'espace de décision au moyen de frontières de séparation. La position relative d'une mesure par rapport à ces frontières aboutit directement à la classification.

De plus deux types de rejet, en ambiguïté et en distance, permettent d'affiner le diagnostic en réduisant le risque de mauvaise classification et en entrouvrant la possibilité de rendre adaptatif le système de décision.

La dernière partie de la thèse consiste à injecter la connaissance de l'expert machine asynchrone dans l'architecture d'un système de décision à base de reconnaissance des formes. En mêlant les spécificités de la machine asynchrone et le caractère générique de l'approche reconnaissance des formes, l'objectif est de proposer un système automatisé de diagnostic de

la machine asynchrone. Au niveau de l'application, sont considérés certains défauts créés au stator et au rotor. Les défauts étudiés sont les cassures de barres au rotor et le déséquilibre d'alimentation. La machine est alimentée à vide et à plusieurs niveaux de charge pour une alimentation prélevée soit sur le réseau, soit par le biais d'un onduleur de tension. Les mesures accessibles en direct sur la machine sont les courants et les tensions électriques.

Deux bibliothèques de mesure ont été constituées :

- l'une dans le cas sans onduleur
- l'autre avec onduleur

Dans un premier temps, une première phase mélangeant connaissances physiques de la machine et traitement de signaux aboutit au calcul d'un nombre conséquent de paramètres (31), candidats au vecteur forme. Les paramètres sont issus d'analyses temporelles et fréquentielles effectuées sur les composantes en quadrature dans les axes  $\alpha, \beta$  et sur les composantes symétriques des tensions et des courants statoriques.

Le nombre élevé de paramètres conduit à une dimension trop importante de l'espace de représentation et éventuellement à une mauvaise définition des observations. De ce fait les algorithmes de sélection sont utilisés pour rechercher les paramètres les plus pertinents et pour réduire le temps de calcul. Deux méthodes de sélection de paramètres sont utilisées afin de choisir le vecteur forme :

- la première utilise le calcul du critère de Fisher pour chaque paramètre,
- la seconde méthode est la procédure appelée "Sequential Backward Selection". Elle calcul un critère traduisant la dispersion des classes sur des sous – ensembles de paramètres.

Un seul et unique vecteur forme (de dimension  $d = 6$ ) est retenu, pour le fonctionnement avec ou sans onduleur. Mais il ne faut pas se tromper d'objectif ; devant une machine en cours de fonctionnement, le type d'alimentation est connu a priori : avec ou sans onduleur. Cela légitime la réalisation de deux bibliothèques. Pour établir un diagnostic de la machine asynchrone alimentée par le réseau, il faut se référer au système de décision appris sur les données "réseau" et non pas onduleur, et vice versa.

Pour pouvoir détecter la présence d'un défaut sur la machine nous avons appliqué deux règles de décision : la règle des  $k - ppv$  et la discrimination linéaire par le calcul direct des

frontières de décision. Ces deux règles de décision donnent de très bons résultats de diagnostic pour les modes de fonctionnements considérés.

Toutefois, la règle des  $k - ppv$  fournit un diagnostic plus cohérent des nouvelles mesures non répertoriées dans l'ensemble d'apprentissage. Ces nouvelles observations sont alors caractérisées par le vecteur forme calculé par la méthode SBS.

Les options de rejet dans les règles de décision jouent un rôle prépondérant dans la détection des nouveaux modes de fonctionnement. Connaissant les signatures propres aux défauts de la machine, il est relativement aisé de caractériser le nouveau mode de fonctionnement associé aux observations rejetées. Elles peuvent alors être intégrées dans l'ensemble d'apprentissage afin d'enrichir la base de données du système de diagnostic.

En conclusion, ce travail a permis de monter l'efficacité de l'application des méthodes de reconnaissance des formes au diagnostic des machines asynchrones. Ici seuls ont été abordé les défauts sur la machine d'entraînement, sans tenir compte de l'apparition par exemple d'un défaut sur l'onduleur et de sa répercussion sur le fonctionnement de la machine. La possibilité d'étendre ces méthodes à une approche globale du diagnostic des systèmes électriques ne fait aucun doute.

Les travaux réalisés dans le cadre de cette thèse pourraient donner lieu à des études supplémentaires sur plusieurs points :

- Normaliser les mesures par rapport au niveau de charge ; au lieu d'avoir pour un mode de fonctionnement donné plusieurs classes correspondant aux différents niveaux de charge.
- Calculer d'autres paramètres avec des méthodes de traitement des signaux plus performantes, pour améliorer la pertinence du vecteur forme. Et rechercher des signatures spécifiques au régime transitoire
- Améliorer les méthodes utilisées en exploitant mieux les rejets de distance pour rendre le système de décision totalement adaptatif, c'est à dire autoriser un nouvel apprentissage du système de décision dès que de nouveaux événements ont modifié la vie du processus.

- Rechercher de nouvelles méthodes, sans modèles en particulier, utilisant les réseaux neuronaux, la logique flou...

---

**BIBLIOGRAPHIE**

---

---

## Contribution Personnelle

### *Conférences Internationales (avec comités de lecture et actes)*

R. Casimir, E. Boutleux, G. Clerc, F. Chappuis, " Comparative study of diagnosis methods for induction motor" Proc. *International Conference on Electrical Machines (ICEM)*, 465-470, 26 - 28 August 2002, Bruges Belgium.

[Casimir, 02] : R. Casimir, E. Boutleux, G. Clerc, F. Chappuis, "A decision system to detect failures in induction motors " Proc. *IEEE International Conference on Systems, Man and Cybernetics*, pp 6, Vol. 5, October 6-9, 2002 Hammamet, Tunisia.

[Casimir, 03 a] : R. Casimir, E. Boutleux, G. Clerc, F. Chappuis, "Broken bars detection in an induction motor by pattern recognition", Proc. *IEEE Power Tech'03 International Conference*, 282 - 288, 23-26 June 2003 Bologna Italy

[Casimir, 03 b] : R. Casimir, E. Boutleux, G. Clerc, " Fault diagnosis in an induction motor by pattern recognition methods", Proc. *IEEE SDEMPED 2003 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, 294 - 299, 24-26 August 2003 Atlanta, GA, USA.

### *Conférence Nationale (avec comités de lecture et actes)*

R. Casimir, "Utilisation de méthodes de reconnaissance des formes pour le diagnostic des machines asynchrones" *Conférence de Jeunes Chercheurs en Génie Electrique (JCGE'03)*, 245 – 250, Saint-Nazaire, 5 et 6 juin 2003.

- 
- [Assaf, 01] T. Assaf, H. Henao, G. A. Capolino, "Detection of voltage source dissymetry using the measurement of symmetrical components in working induction motors", *IEEE SDEMPED'01 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives* pp 441 – 447, 1 – 3 September 2001, Grado Italy
- [Benbouzid, 99] M. E. H. Benbouzid, M. Vieira, C. Theys, "Induction motor's faults detection and localization using stator current advanced signal processing techniques" *IEEE Transaction on Power Electronics*, Vol. 14, N° 1, pp 14 – 22, January 1999
- [Boumégoura, 99] T. Boumégoura, H. Yahoui, G. Clerc, G. Grellet, "Observation des paramètres du moteur asynchrone à cage d'écuréuil avec un observateur non linéaire" *Colloque EF'99* - pp 375 – 379 Lille 30&31 mars 1999.
- [Boutleux, 96] E. Boutleux, "Diagnostic et suivi d'évolution de l'état d'un système, par reconnaissance des formes floues. Application au modèle du réseau téléphonique français" *Thèse de doctorat Université de technologie de Compiègne 1996*
- [Cabanas, 98] H. F. Cabanas, M. Garcia-Melero, "Experimental applications of axial leakage flux to the detection of rotor asymmetries, mechanical anomalies and interturn short-circuits in working induction motors" *Proc ICEM'98*, Vol 1/3, September 2 -4, 1998, Istanbul Turkey.
- [Calis, 99] H. Calis, P. J. Unsworth, "Fault diagnosis in induction motors by motor current signal analysis" *Proc. IEEE SDEMPED'99 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 237 – 241, September 1 –3, Gijon, Spain.
- [Casimir, 02] R. Casimir, E. Boutleux, G. Clerc and F. Chappuis, "A decision system to detect failures in induction motors". *Proc IEEE International Conference on Systems, Man and Cybernetics* Vol 5, October 6-9, 2002 Hammamet, Tunisia
- [Casimir, 03 a] R. Casimir, E. Boutleux, G. Clerc, F. Chappuis, "Broken bars detection in an induction motor by pattern recognition", *Proc. IEEE Power Tech'03*

---

*International Conference*, pp 282 - 288, 23-26 June 2003, Bologna Italy

- [Casimir, 03 b] R. Casimir, E. Boutleux, G. Clerc, " Fault diagnosis in an induction motor by pattern recognition methods", *IEEE SDEMPED 2003 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 294 - 299, 24-26 August 2003 Atlanta, GA, USA
- [Chen, 03]. X. Chen, "An improved branch and bound algorithm for feature selection". *Pattern Recognition letter* Vol 24 pp 1925 – 1933, 2003
- [Chouiter, 97] D. R. Chouiter, "Conception et réalisation d'une commande robuste de machine asynchrone", *Thèse de doctorat école centrale de Lyon janvier 1997*.
- [Deng, 93] X. Deng, E. Ritchie, "Detection of broken bars in induction motors by investigating the flux linkages of the stator winding" *ICEMA 14 - 16* September 1993 Adelaide Australia.
- [Devijver, 77] P. A. Devijver, "Reconnaissance des formes par la méthode des plus proches voisins", *Thèse de 3<sup>ème</sup> cycle, Université Paris 6, juin 1977*
- [Diday, 76] E. Diday, J. C. Simon, "*Cluster Analysis*", dans *Digital Pattern Recognition*, (K. S. FU edition), pp 47 – 94, Springer – Verlag, Berlin, 1976
- [Didelet, 92] E. Didelet, "Les arbres de neurones avec rejet d'ambiguïté. Application au diagnostic pour le pilotage en temps réel du réseau téléphonique français" *Thèse de doctorat Université de technologie de Compiègne 1992*
- [Dubuisson, 90] B. Dubuisson, *Diagnostic et reconnaissance des formes*. Traité des nouvelles technologies. Série diagnostic et Maintenance. Hermès, Paris, 1990
- [Dubuisson et al., 01] B. Dubuisson (sous la direction de ~) , *Diagnostic, intelligence artificielle et reconnaissance des formes*. Traité IC2 (Informatique - Commande - Communication). Série Productique. Hermès, Paris 2001.
- [Duda, 73] R. O. Duda, P. E. Hart, "*Pattern Classification and Scene Analysis*", John Wiley and Sons, New York, 1973.



- 
- [Fiser, 96] R. Fiser, S. Ferkolj, "Modelling of failure states of induction machines" *Electrotechnical conference, MELECON'96 8<sup>th</sup> Mediterranean*, vol 3 pp 1195 - 1198.
- [Fisher, 36] R. A. Fisher, "The use of multiple measurements in taxonomic problems". *Annals of Eugenics*, v. 7, 1936, pp 179 – 188
- [Friedman, 99]. M. Friedman and A. Kandel, *Introduction to pattern recognition – statistical, structural, neural and fuzzy logic approaches*. Imperial College Press, London 1999.
- [Fukunaga, 72] K. Fukunaga, *Introduction to statistical pattern recognition*. Academic Press, New York, 1972
- [Fukunaga, 84] K. Fukunaga and T. E. Flick, "An optimal global nearest neighbor metric". *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, no. 3, may 1984
- [Govaert, 03] G. Govaert (Sous la dir. De), *Analyse des données*. Traité IC2 – Information, Commande, Communication. Série traitement du signal et de l'image. Hermès, Paris 2003
- [Hellman, 70] M. E. Hellman, "The nearest neighbor classification rule with a reject option", *IEEE. Trans on System, Man and Cybernetics*, pp 179 – 185, Vol 6, no 3, July 1970.
- [Hénao, 99] H. Hénao, M. Garcia-Melero, "Méthodes de détection de défauts dans les machines à induction par utilisation de la mesure du flux axial" *Colloque EF'99 - Lille, 30 et 31 mars 1999 pp 106 - 111*
- [Koford, 66] J. S. Koford, G. F. Groner, "The use of an adaptive threshold element to design a linear optimal pattern classifier", *IEEE Trans on Information Theory*, pp 42 – 50, Vol. 12, no 1, January 1966
- [Kral, 99] C. Kral, F. Pirker, "Rotor eccentricity detection of induction machines by torque estimation" *Proc. IEEE SDEMPED'99 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 283 - 287,

---

September 1 –3, Gijon, Spain.

- [Kudo, 00] M. Kudo, J. Sklansky, "Comparison of algorithms to select features for pattern classifiers". *Pattern Recognition* Vol. 33, Issue 1, pp 25-41 January 2000
- [Lara, 95] S. C. De Lara Jayme, "Identification des paramètres des machines électriques à courant alternatifs" *Thèse de doctorat INP Toulouse 1995*.
- [Moreau, 99] S. Moreau, J. C. Trigeassou, G. Champenois, "Diagnosis of electrical machines : a procedure for electrical fault detection and localization" *Proc. IEEE SDEMPED'99 – Symposium on Diagnostics for Electric Machines, Power Electronics and Drives*, pp 225 - 229, September 1 –3, Gijon, Spain.
- [O'Donnell, 85] P. O'Donnell, "Report of large reliability survey of industrial and commercial installations, Part 1", *IEEE Transaction on Industry Applications*, Vol. IA 21, N° 4, july/august 1985 pp 853 - 864
- [Ouaghram, 99] R. B. Ouaghram, "Contribution au diagnostic des machines tournantes complexes. Application à un laminoir", *Thèse de doctorat Université de technologie de Compiègne 1999*
- [Pinkowski, 97] B. Pinkowski, "Principal Component analysis of speech spectrogram images". *Pattern recognition*, Vol. 30, no. 5, pp. 777-787, 1997.
- [Saïd, 00] M. S. N. Saïd, M. E. H. Benbouzid et A. Benchaib, "Detection of broken bars in induction motors using an extended kalman filter for rotor resistance sensorless estimation," *IEEE Transactions on Energy Conversion*, vol 15, no 1, mars 2000
- [Schaeffer, 99 a] E. Schaeffer, "Diagnostic des machines asynchrones : modèles et outils paramétriques dédiés à la simulation et à la détection de défauts" *Thèse de doctorat Ecole Centrale de Nantes 1999*
- [Schaeffer, 99 b] E. Schaeffer, E. Le Carpentier, E. H. Zaïm, L. Loron, "Diagnostic des entraînements électriques : détection de courts-circuits statoriques dans la machine asynchrone par identification paramétrique" *17<sup>ième</sup> colloque*

---

*GRETSI sur le traitement du signal et des images*, Vol 4, pp 1037 - 1040, Vannes, France, 13 - 17 septembre 1999.

- [Schoen, 95 a] R. R. Schoen, B. K. Lin, T. G. Habetler, "An unsupervised, on-line induction motor fault detection using stator current monitoring" *IEEE Transaction on Industry Applications*, Vol. 31, N° 6, november/december 1995 pp 1280 - 1286
- [Schoen, 95 b] R. R. Schoen, T. G. Habetler, F. Kamran, R. G. Bartheld, "Motor bearing damage detection using stator current monitoring" *IEEE Transaction on Industry Applications*, Vol. 31, N° 6, november/december 1995 pp 1274 - 1279
- [Sobczyk, 98] T. J. Sobczyk, A. Izvorski, "Recognition of rotor eccentricity of induction motor based on the fourier spectra of phase currents" *Proc. ICEM'98*, pp 408 - 413 , Vol 1, september 2 - 4 1998 Istanbul Turkey
- [Somol, 00] P. Somol, P. Pudil, F.J. Ferri and J. Kittler, "Fast branch and bound algorithm in feature selection". *Proc SCI / ISAS, Vol 7 World Multiconference on Systemics and Informatics (SCI) 2000*, pp 646-651
- [Sonbaty] Y. E. Sonbaty, M. A. Ismail, "On-line hierarchical clustering", *Pattern Recognition letter* Vol 19 pp 1285 - 1291, 1998
- [Stein, 88] J. Stein, R. D. Endicott, "Noninvasive detection of broken rotor bars in operating induction motors" *IEEE Transactions on Energy Conversion*, Vol 3, N° 4, december (1988)
- [Thollon, 93] F. Thollon, G. Grellet, A. Jammal, "Asynchronous motor cage fault detection through electromagnetic torque measurement" *Proc ETEP* Vol 3, N° 3, September/October 1993
- [Thomson, 00] Thomson, W.T.; Fenger, M. "Industrial application of current signature analysis to diagnose faults in 3-phase squirrel cage induction motors " *Pulp and Paper Industry Technical Conference, 2000*. Conference Record of 2000, pp 205 -211

- 
- [Tomek, 76] I. Tomek, "A generalisation of the kNN rule", *IEEE Trans on System, Man and Cybernetics*, Vol. 6, no 2, pp 121 – 126, february 1976
- [Vas, 92] P. Vas, "Electrical machines and drives, a space vector theory approach" Oxford University Press, New York, USA 1992
- [Wilson, 97] D. R. Wilson, T. R. Martinez, "Improved heterogeneous distance functions", *Journal of Artificial Intelligence Research*, Vol 6, pp 1 – 34, (1997)
- [Yahoui, 95] H. Yahoui, "Contribution à la surveillance et au diagnostic des machines électriques. Application aux machines triphasées asynchrones à cage." *Thèse de doctorat - Université Claude Bernard - Lyon1 1995.*
- [Yahoui, 96] H. Yahoui, J. Seetohul, G. Grellet, "Detection of broken bar or ring in asynchronous machines by the spectrum analysis of the observed electromagnetic torque" Electricmacs'96, Saint Nazaire (France), *septembre 1996.*
- [Zieba, 95] S. Zieba, "Une méthode de suivi d'un système évolutif. Application au diagnostic de la qualité d'usinage" *Thèse de doctorat Université de technologie de Compiègne 1995*
- [Zwingelstein, 95] G. Zwingelstein, "Diagnostic des défaillances. Théorie et pratique pour les systèmes industriels" *Traité des nouvelles technologies - série Diagnostic et Maintenance 1995. Editions HERMES.*

**Titre :****Diagnostic des Défauts des Machines Asynchrones par Reconnaissance Des Formes****Mots clés :**

Diagnostic – Reconnaissance des formes – Machines asynchrones – Défauts – Classification – Détection – Décision.

**Résumé :**

Les progrès réalisés en matière d'électronique de puissance, de circuits de commande et en automatique ont contribué à l'utilisation grandissante des machines asynchrones dans les systèmes d'entraînements électriques. Le recours aux machines asynchrones est surtout lié à leur robustesse, leur puissance massique et à leur coût de fabrication. Leur maintenance et leur surveillance permettent de rentabiliser les installations. Il est donc important de développer des outils de diagnostic pour détecter de manière précoce les défauts pouvant apparaître dans ces machines.

Notre approche est basée sur l'utilisation des méthodes de reconnaissance des formes. Un vecteur de paramètres, appelé vecteur forme, est extrait de chacune des mesures effectuées sur la machine. Les règles de décisions utilisées permettent de classer les observations, décrites par le vecteur forme, par rapport aux différents modes de fonctionnement connus avec ou sans défaut.

Des défauts ont été créés au rotor et au stator de la machine asynchrone, alimentée soit à partir du réseau, soit par le biais d'un onduleur de tension.

Les procédures de décisions, basées sur la règle des k - plus proches voisins et sur le calcul direct des frontières, ont été utilisées pour détecter les défauts. Ces algorithmes ont montré l'efficacité de l'application de la reconnaissance des formes au diagnostic.

***Title: Faults Diagnosis of Induction Motors by Pattern Recognition.***

***Keywords: Diagnosis – Pattern Recognition – Induction Machines – Faults – Classification – Detection – Decision.***

---

**Direction de Recherche :**

Monsieur Guy CLERC, Professeur des Universités  
Centre de Génie Electrique de Lyon (CEGELY) – UMR CNRS 5005  
Université Claude Bernard Lyon 1, bât. Omega  
43 Boulevard du 11 novembre 1918  
69622 Villeurbanne (France)