Fault Detection in Nonlinear Systems

by

Adam A. Papadimitropoulos,

Thesis

Presented to the Department of Electronic Engineering of

Technical University of Crete, GREECE

in Partial Fulfillment

of the Requirements

for the Degree of

Master in Electronic Engineering

Technical University of Crete, GREECE

August 2003

To my parents, Andreas and Cleo

Alla famiglia di Pica

Acknowledgments

A long list of friend and colleagues have helped me in the preparation of the thesis in many different ways. I would like to express my gratitude to the professors George Rovithakis and Thomas Parisini for their continuous guidance, encouragement and support. Without their insightful discussions and contribution, this thesis would have never been written. They have both introduced me in the fascinating world of Control and Fault Diagnosis as well as they have sharpened my skills towards these fields. Their outstanding academic career and as persons themselves, they were and always will be a source of inspiration for me.

Special thanks to professor M. Christodoulou for accepting me as a postgraduate student in technical university of Crete and for his cooperation in the preparation of this thesis and his useful suggestions in finetuning my research.

Thanks are due to the members of DAMADICS research training network for accepting me as a member and for the finance of my work as long as I stayed in Italy.

I would like to thank my parents and my brother Harry, for all the encouragement and support they have given to me my entire life. I would not be where I am today without them. Special thanks also to my lifelong friends Sifis and Papadopoulos. At the end, I would like to thank Stefania and her family as well as Parisini's family,

for their love, endurance and the true emotions that were supporting me during the tough years of my graduate studies in Italy. I would like to express my deepest and most frank appreciation to them and especially to Stefania.

ADAM A. PAPADIMITROPOULOS

Technical University of Crete, GREECE

August 2003

Absrtact

In this thesis, various non smooth nonlinearities presented in any physical system and fault diagnosis methods are examined. Towards this concept, a large number of mathematical models and their identification and estimation techniques are presented. In parallel, an introduction in the fault diagnosis area and its up-to-date methodology are also presented.

The problem of actuator fault detection in mechanical systems with friction that perform linear motion, is discussed and it is the main contribution of this dissertation. The dynamic LuGre model is used to model the effects of friction. The proposed architecture is built upon an on-line neural network approximator which requires only system's position and velocity. The friction internal state is not assumed to be available for measurement. The developed fault detector is analyzed with respect to its robustness and sensitivity. Rigorous fault detectability conditions and upper bounds for the detection time are also derived. The proposed methodology is applied to the DAMADICS benchmark problem which is developed in order to approximate the industrial process in a sugar factory located in Lublin (Poland). The neural network approximation scheme makes it possible to detect either incipient or abrupt faults regarding the friction and the spring models of the considered actuator.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The need to design systems able to guarantee increased reliability, availability and safety, triggered on-going research in the area of fault diagnosis, mainly through the model-based analytical redundancy path. Analytical redundancy schemes have received considerable attention in the last two decades mostly owing to the advances in computer technology, as well as to the appearance of powerful signal processing and learning methodologies. In general, the actual behavior of the plant is compared to that expected, on the basis of a plant model. Deviations between the actual and the estimated behavior are expressed in terms of residuals, which are indications of faults. Detailed overviews of such schemes may be found in [26]-[29]. The majority of these methods are constraint to linear systems. Owing to the inherent complexity, derivation of analytical results regarding robustness and sensitivity of fault diagnosis schemes for nonlinear systems is difficult. Despite the difficulties, works on nonlinear systems have recently appeared [51]-[53] and [54].

In a wide range of physical systems such as mechanical systems, electro-magnetic systems, actuators, sensors etc., non-smooth nonlinear mechanisms such as friction, backlash and hysteresis, severely limit their performance and reliability. Up to now,

previous works on nonlinear fault diagnosis were focused on developing general-purpose architectures. Thus, unavoidably, they were restricted to special classes of nonlinear systems, assuming full state measurement and smooth nonlinearities. Nonlinear observers have also been employed to relax the full state measurement assumption. Unfortunately, the use of observers further restricts the class of nonlinear systems and the type of permissible faults. Moreover, nonlinear observer design is not de-coupled from controller design. Hence, important theoretical questions regarding even fault detection are raised, since practically all systems operate in a closed loop.

The main contribution of this thesis is that of detecting faults in mechanical systems with friction. Friction is present in any system that involves mechanical motion. It may cause large steady state errors and oscillations generated by a combination of friction, which counteracts motion, and an instability mechanism, thus making friction a very complicated phenomenon. The aforementioned reasons impose extra complexity to any scheme that is targeted at diagnosing faults in such systems.

Studies [3], [6] have shown that a friction model involving dynamics is necessary to describe accurately the friction phenomena. Various dynamic models have been proposed [9], [3], [15] and [18]. However, the unknown structure of the incoming faults significantly magnifies the level of system uncertainty. Neural networks with their massive parallelism, very fast adaptability and inherent approximation capabilities, have already been utilized mainly towards the friction compensation problem [10].

In this work we present a novel approach to detect faults in mechanical systems with friction that perform linear motion. The basic module in the proposed ar-

chitecture is an on-line approximator which is based on liner-in-the-weights neural network structures. To model the effects of friction, the dynamic LuGre model [9] is used. However, we don't assume knowledge of system nonlinearities. Furthermore, the friction internal state is not assumed to be available for measurement. The on-line approximator requires system's position and velocity as well as its input force. The performance of the developed fault detector is analyzed with respect to its robustness and sensitivity. Rigorous fault detectability conditions are also derived basing on the important results presented in [51]. We go beyond the theoretical analysis and present simulation studies to clarify and verify the approach with emphasis on the application to the DAMADICS actuator benchmark problem. Under the framework of the DAMADICS research network funded by the European Union, a benchmark model was developed to approximate the behavior of the evaporation stage of a sugar factory in Lublin (Poland). Actuators under consideration consist of a control valve, a pneumatic linear servomotor and a positioner. In such kind of electromechanical systems, the presence of friction phenomena is unavoidable and significantly increases the complexity of the fault diagnosis problem.

The thesis is organized as follows. In Chapter 2, wide-used models in the literature for friction, backlash and hysteresis are presented. It is also reported their limitations and applicability. Additionaly the existed identification and estimation techniques are also investigated. The main purpose of this chapter is to explain in-depth the non-smooth non-linearities that are often appear to the physical systems. Furthermore, as it is explained above, the model constitutes the basis for the development of model-based analytical redundancy methods. In Chapter 3 an introduction to the fault diagnosis area and some basic definitions are being made. This chapter presents some basic methods, such as observers, parity relations, etc. It also presents the up-to-date research where it is oriented to the nonlinear sys-

tems and the application of qualitative and computational intelligence techniques. In Chapter 4, which is a combination of the two preceded chapters, our method is presented. We formulate the problem and we state the necessary assumptions. Some definitions and preliminaries are also provided. This chapter also deals with the design and the robustness analysis of the on-line approximation scheme. Moreover, the sensitivity analysis of the fault detection scheme is carried out, in which fault detectability conditions are also derived. In addition, upper bounds on the detection time and a relationship between detection time and the values of certain design parameters are established. Simulation studies are also presented. Finally, in Chapter 5, a brief description of the DAMADICS benchmark problem including its simulation results and the performance of the proposed, in Chapter 4, methodology are given. The results are clarify and verify the reliability of the proposed method.

Chapter 2

Non-smooth non-linearities

Non-smooth non-linearities are common in practical systems. Such non-linearities are usually poorly known and may vary with time and they often severly limit system performance. Especially in actuators which are installed in harsh environment, non-linearities increase with wear and tear and in mass production change from component to component. The objective is to design a desirable system in order to be able to accomodate such uncertainties. Typical non-smooth non-linearities addressed in this chapter are friction, backlash and hysteresis. In the following, various models used in the literature and some approximation techniques regarding friction, backlash and hysteresis phenomena are presented.

2.1 Friction

Whenever there is a motion or tendency of motion between two elements, friction forces exist. The frictional forces encountered in physical systems are usually of a non-linear nature. The characteristics of the frictional forces between the surfaces often depend on such factors as the composition of the surfaces the pressure between the surfaces, their relative velocity and others, so that an exact mathematical

description of the frictional force is difficult to be established.

The types of friction which are commonly used in practical systems are : **Viscous**, **Static**, **Coulomb**, and **Stribeck** friction. The friction model being used is a conglomerate of these friction components from which a force balance may be obtained for friction acting against a surface [5]. It is often assumed when studying friction that there is no motion while in static friction, which is to say no motion without sliding. But Dahl [15], [16], [17] studying experimental observation of friction in small rotation of ball bearing concluded that for small motions, a junction in static friction behaves like a spring and considered the implications for control. There is a displacement (pre-sliding displacement) which is an approximately linear function of the applied force, up to a critical force, at which breakaway occurs. When forces are applied, the asperities will deform, but recover when the force is removed. At this point, the tangential force is governed by:

$$F_t(x) = -k_t x \tag{2.1}$$

where F_t is the tangential force, k_t is the tangential stiffness of the contact and x is the displacement away from the equilibrium position. F_t and x refer to the force and displacement in the contact before sliding begins. The tangential stiffness k_t, is a function of asperity geometry, material elasticity and applied normal force. To first approximation it is actually the breakaway displacement that is constant and the stiffness is then given by

$$k_t = \frac{F_b}{x_b} \tag{2.2}$$

where F_b is the breakaway force and x_b the maximum deformation. The transition from elastic contact to sliding is not simple. Sliding is observed to originate first at the boundary of a contact and to propagate toward the center. Thus there is no abrupt transition to sliding. Pre-sliding displacement is of interest to the control community in extremely high precision pointing applications in dynamics and in simulation and may also be important in establishing that there are no

discontinuities in friction as a function of time.

Coulomb friction from the other hand, is a retarding force that has a constant amplitude with respect to the change in velocity, but the sign of the frictional force changes with the reversal of the direction of velocity. The other friction component, Viscous friction, represents a retarding force that is a linear relationship between the applied force and velocity. Both of these phenomena is obviously that opposing the motion when the velocity is different from zero.

However imperfection in the motor mechanics and unbalances on the motor shaft yield asymmetries behavior of the motor dynamics. The model proposed by [7] includes Coulomb and viscous friction and accounts for friction asymmetries. Some experiments conducted with this model showed that the asymmetries in the Coulomb friction components were dominant. This observation was also corroloborated by the results presented in [2].

An another effect takes place after the stiction force has been surmounted where the friction force decreases exponentially reaching approximately 60% of the breakaway force ([8]), and then increases proportionally to the velocity. These bends occur at velocities close to zero. This type of friction structure, sometimes known as a stick-slip friction. This arises because static friction is greater than the level of Coulomb friction at zero velocity. Stribeck friction can be explained as an inertial effect occuring when trying to separate two objects which have been at rest for long periods of time. The Stribeck friction force decreases as movement occurs. The phenomena of friction decreasing during a sliding period after movement is called stick-slip. To capture this behavior an empirical Stribeck velocity parameter, the so-called v_s, is used as we shall show later in the following mathematical models of friction. The Cincinnati Milacron test procedure [14] indicates that when $F_s/F_c < 0.85$ stick-slip will be eliminated. It is also widely observed that stick-slip can be eliminated by stiffening a mechanism. The following expression summarize the main friction

components:

$$F_f(v) = asgn(v) \tag{2.3}$$

$$F_f(v) = a_i sgn(v) + b_i v \tag{2.4}$$

$$F_f(v) = (a_0 + a_1 e^{-b|v|}) sgn(v) \tag{2.5}$$

where a in (2.3) represents the Coulomb friction, a_i, b_i in (2.4) the asymmetric model of Coulomb and viscous friction. In (2.5) the sum represents the breakaway force and b the slip constant.

We note, before presenting the dominant models, that when the velocity is not constant, the dynamics of the model will be very important and give rise to different types of phenomena such as friction lag (frictional memory); a change in friction will lag changes in velocity or load. Also the friction force is lower for decreasing velocities than for increasing velocities responding to the existence of the hysteresis in the relation between friction and velocity. The hysteresis loop becomes wider at higher rates of the velocity changes. Hess & Soom explained their experimental results by a pure time delay in the relation between velocity and friction force.

Finally an another time-dependent property of friction is the rising static friction with increasing dwell time. Dwell time is the time spent in static friction. [25] proposed an empirical model that incorporates the relation between static friction and dwell time as:

$$F_s(t) = F_{s,\infty} - (F_{s,\infty} - F_c)e^{-\gamma t^m} \tag{2.6}$$

where $F_{s,\infty}$ is the ultimate static friction; F_c the Coulomb friction at the moment of arrival in the stuck condition; γ, m are empirical parameters. [3], [4] presents a model of rising static friction which is useful for analysis and solves some problems associated with using F_c as the starting point of the static friction rise. The model is:

$$F_{s,b_n}(t_2) = F_{s,a_{n-1}} + (F_{s,\infty} - F_{s,a_{n-1}})\frac{t_2}{t_2 + \gamma} \tag{2.7}$$

8

where F_{s,b_n} is the level of Stribeck friction at the beginning (breakaway) of the nth interval of slip; and $F_{s,a_{n-1}}$ is the Stribeck friction at the end (arrival) of the previous interval slip. Note γ is still an empirical factor, will be different in physical dimension from the equation (2.6).

2.1.1 Mathematical Models of Friction

In [3], proposed a seven parameter model, where the friction is given by:

- Not sliding (presliding displacement)

$$F_f(x) = -k_t x \tag{2.8}$$

- Sliding (Coulomb + viscous + Stribeck curve function with frictional memory)

$$F_f(\dot{x}, t) = -(F_c + F_v|\dot{x}| + F_s(\gamma, t_2)\frac{1}{1 + (\frac{\dot{x}(t-\tau_L)}{\dot{x}_s})^2})sgn(\dot{x}) \tag{2.9}$$

- Rising static friction (friction level at breakaway)

$$F_s(\gamma, t_2) = F_{s,a} + (F_{s,\infty} - F_{s,a})\frac{t_2}{t_2 + \gamma} \tag{2.10}$$

where :

- $F_f()$ is the instantaneous friction force

- F_c (*) is the Coulomb friction force

- F_v (*) is the viscous friction force

- F_s is the magnitude of the Stribeck friction (frictional force at breakaway is $F_c + F_s$)

9

- $F_{s,a}$ is the magnitude of the Stribeck friction at the end of of the previous sliding period

- $F_{s,\infty}$ (*) is the magnitude of the Stribeck friction after a long time at rest (with a slow application of force)

- k_t (*) is the tangential stifness of the static contact

- $\dot{x}_s$ (*)is the characteristic velocity of the Stribeck friction

- τ_L (*) is the time constant of frictional memory

- γ (*) is the temporal parameter of the rising static friction

- t_2 is the dwell time, time at zero velocity.

(*) marks friction model parameters, other variables are state variables.

The magnitude of the seven friction parameters will naturally depend upon the mechanism and lubrication, but typical parameters may be offered. These are summarized in Table 2.1 and in Table 2.2 is presented for each of seven parameters of the model where represent a different friction phenomenon, the effect of these phenomena on sliding behavior. This model, however, does not combine the different friction phenomena but it is in fact one model for stiction and another for sliding friction. Another dynamic model suggested by Rice and Ruina has been used in connection with control by Dupont. This model is not defined at zero velocity.

In [8] a friction model covering most of friction components can be expressed as follows:

$$F_f(\dot{x}) = [a_0 + a_1 e^{-b_1|\dot{x}|} + a_2(1 - e^{-b_2|\dot{x}|})]sgn(\dot{x}) \qquad (2.11)$$

where a_i's and b_i's are positive constants. Asymmetries can be included in (2.11) by letting a_i's be different for different velocity direction. The b_i's can be maintained

Parameter	Range	Parameters depends principally upon
F_c	$0.001 - 0.1 * F_n$	Lubricant viscosity, contact geometry and loading
F_v	0-very large	Lubricant viscosity, contact geometry and loading
$F_{s,\infty}$	$0 - 0.1 * F_n$	Boundary lubrication,F_c
k_t	$\frac{1}{\Delta x} * (F_s + F_c); \Delta x \simeq 1 - 50[\mu m]$	Material properties and surface finish
$\dot{x}_s$	$0.00001 - 0.1[\frac{meter}{second}]$	Bound. lubric., lubricant viscosity, contact geometry and loading
τ_L	$1 - 50[ms]$	Lubricant viscosity, contact geometry and loading
γ	$0 - 206[s]$	Boundary lubrication

Table 2.1: Approximate ranges for the parameters of seven parameter friction model

Friction model	Predicted/Observed behavior
Viscous	Stability at all velocities an at velocities reversal
Coulomb	No stick-slip for Pd control;No hunting for PID control
Static+Coulomb+Viscous	Predicts stick-slip for certain initial conditions under PD control;predicts hunting under PID control
Stribeck	Needed to correctly predict initial conditions leading to stick-slip
Rising static friction	Needed to correctly predict interaction of velocity and stick-slip amplitude
Frictional memory	Needed to correctly predict interaction of stiffness and stick-slip amplitude
Presliding	Needed to correctly predict small displacements while sticking (including velocity reversals)

Table 2.2: Friction model capabilities

constant. The non-linearity of parameters a_i and b_i restricts its utility for on-line identification (linear predictors require a model expression that it is linear in parameters). A simplified model is also presented in the same paper that captures the asymmetries and stick-slip while remains linear in the unknown parameters. Such model is :

$$F_f(\dot{x}) = [a_0 + a_1|\dot{x}|^{\frac{1}{2}} + a_2|\dot{x}|]sgn(\dot{x}) \qquad (2.12)$$

To evaluate the precision that can be achieved with this reduction the parameters a_i of the model (2.12) are estimated by minimizing a least-square estimation algorithm. In this paper also referred that the uniqueness of the a_i's does not exist. Indeed, several sets of parameters a_i may exist leading to equivalent approximation.

The LuGre model, its variants and approximation techniques

In [9] the friction interface between two surfaces is presented in some extent, as a contact between bristles. The average deflection of the bristles is denoted by z and is modelled by:

$$\frac{dz}{dt} = v - \frac{|v|}{g(v)}z \qquad (2.13)$$

where v is the relative velocity between the two surfaces. The first term gives a deflection which is proportional to the integral of the relative velocity. The second term asserts that the deflection z approaches the value:

$$z_{ss} = \frac{v}{|v|}g(v) = g(v)sgn(v) \qquad (2.14)$$

in steady-state, i.e, when v is constant. The function g is positive and depends on many factors such as material properties, lubrication, temperature. It needs not be symmetrical. Direction dependent behavior can therefore be captured. For typical bearing friction, $g(v)$ will decrease monotonically from $g(0)$ when v increases. This corresponds to Stribeck effect. The friction force generated from the bending of the

bristles is described as:

$$F = \sigma_0 z + \sigma_1 \frac{dz}{dt} + \sigma_2 v \qquad (2.15)$$

where σ_0 is the stiffness, σ_1 a damping coefficient and $\sigma_2 v$ a term which accounts for viscous friction. The function $\sigma_0 g(v)$ and $\sigma_2 v$ can be determined by measuring the steady-state friction force when the velocity is held constant. A parameterization of g that has been proposed to describe the relation between velocity and friction force for steady-state motion is given by:

$$
\begin{aligned}
F_{ss}(v) &= \sigma_0 g(v) sgn(v) + \sigma_2 v & (2.16) \\
&= F_c sgn(v) + (F_s - F_c) e^{-(\frac{v}{v_s})^2} sgn(v) + \sigma_2 v & (2.17)
\end{aligned}
$$

In that paper is assumed that if the parameters $\sigma_0, \sigma_1, \sigma_2$ and function $g(v)$ are known and using a non-linear friction observer to estimate the unmeasurable state z where the observer given by:

$$
\begin{aligned}
\frac{d\hat{z}}{dt} &= v - \frac{|v|}{g(v)}\hat{z} - Ke & (2.18) \\
\hat{F} &= \sigma_0 \hat{z} + \sigma_1 \frac{d\hat{z}}{dt} + \sigma_2 v & (2.19)
\end{aligned}
$$

where $K > 0$ and e is the position error, we can have position control. If x_d is the desired reference and is assumed to be twice differentiable then the position error defined as $e = x - x_d$ and the term Ke in (2.18) 'is a correction term for the position error.

Similarly is proposed the velocity control where $e = v - v_d$ with v_d the desired velocity which is assumed to be differentiable.

By the way, to assume that the friction model and its parameters are known exactly is of course a strong assumption. In addition to this the accuracy required in the velocity measurement is a similar problem. The model of friction given by (2.17) is used by [1] where the $sgn()$ function approximated by $tanh(\sigma x)$ function where σ defines the slope of the function. The larger the value the steepest the slope

is. They suggest that a value of $\sigma = 30$ provides a close fit while capturing most of friction effects. They also suggest that the parameters may be estimated, as a structured disturbance, using an observer. A new system then is created so that the parameters become additional states augmented to the original state-space system. Using the non-linear Luenberger observer for parameter estimation, can also be served as an on-line method for detecting faults. However, an extra attention needs to be paid for the parameter selection for estimation where the observability needs to be maintained. In some applications the friction variation may also depend on the actual value of the position or a more complex combination of the position and velocity. As a consequence, in some applications it will be required that friction is explicitly parameterized not only as a function of velocity but also as a function of position (see [10]). Another form of the model described previously in [10] is :

$$\dot{z} = -\alpha(\dot{x})|\dot{x}|z + \dot{x} \tag{2.20}$$

$$F = \sigma_0 z + \sigma_1 \dot{z} + \sigma_2 \dot{x} \tag{2.21}$$

where z denotes the average deflection of the bristles, which is not measurable, $a(\dot{x})$ a finite positive function. One parameterization of $\alpha(\dot{x})$ which describes the Stribeck effect is :

$$\alpha(\dot{x}) = \frac{\sigma_0}{f_c + (f_s - f_c)e^{-(\frac{\dot{x}}{\dot{x}_s})^2}} \tag{2.22}$$

where f_c the Coulomb friction level, f_s is the level of stiction force and $\dot{x}_s$ the Stribeck velocity. The parameters σ_0, σ_1 are assumed to be known. In this case, $\sigma_0/f_c \leq \alpha(\dot{x}) \leq \sigma_0/f_s$, if it is assumed that $f_s \geq f_c$. As it mentioned above the friction may be position dependent. In [10] assumes in this case that $\alpha(x, \dot{x})$ is an upper and lower bounded positive smooth function of x and $\dot{x}$. There is no need to know the exact form of the function as generalized basis functions shall be used to emulate it. In this way it can capture properties related not only to velocity but

also to position. So the model can be:

$$\dot{z} = -\alpha(x, \dot{x})|\dot{x}|z + \dot{x} \tag{2.23}$$

$$F = \sigma_0 z + \sigma_1 \dot{z} + \sigma_2 \dot{x} \tag{2.24}$$

If $\alpha(x, \dot{x})$ is assumed completely unknown neural networks is a possible tool to approximate the non-linear mapping. The approximation $\alpha_\alpha(W, x, \dot{x})$ is written as:

$$\alpha_\alpha(x, \dot{x}) = W^T S(x, \dot{x}) \tag{2.25}$$

where $W = [w_1, w_2, \ldots, w_l]^T \in \Re^l$ the parameter vector and $S(x, \dot{x}) = [s_1(x, \dot{x}),$ $s_2(x, \dot{x}), \ldots, s_l(x, \dot{x})]^T \in \Re^l$ is the vector bounded basis functions, and therefore we have :

$$\alpha(x, \dot{x}) = W^T S(x, \dot{x}) + \epsilon \tag{2.26}$$

with ϵ being the modelling error which is assumed to be bounded. If the non-linear function $\alpha(x, \dot{x})$ is in the functional range of the approximation, then $\epsilon = 0$. For the case where $\alpha(\dot{x})$ is only a function of $\dot{x}$, the rough form/shape of $\alpha(\dot{x})$ in terms of velocity is infinitely smooth. This piece of information, as they suggested, helps to find more appropriated basis function for approximation rather than constructing NN blindly as in most cases. The function is :

$$\alpha(\dot{x}) = \frac{c_0}{1 + c_1 e^{-(\dot{x}/\dot{x}_s)^2}} \tag{2.27}$$

where $c_0 = \sigma_0/f_c$, $c_1 = (f_s - f_c)/f_c$ and $c_2 = \dot{x}_s^2$. The parameters $c_1 > 0$ and $c_2 < 1$ appear non-linearly. The polynomial approximation which is suggested is to expand (2.27) using Taylor expansion around $\dot{x}^2 = 0$. Then :

$$\alpha(\dot{x}) = \sum_{k=0}^{m} \frac{1}{k!} \frac{\vartheta^k \alpha}{\vartheta(\dot{x}^2)^k}|_{\dot{x}^2=0}(\dot{x}^2)^k + \epsilon_t = W^T S(\dot{x}) + \epsilon_t \tag{2.28}$$

where :

$$W = [\alpha|_{\dot{x}^2=0}, \frac{\vartheta \alpha}{\vartheta \dot{x}^2}|_{\dot{x}^2=0}, \ldots, \frac{1}{m!} \frac{\vartheta^m \alpha}{\vartheta(\dot{x}^2)^m}|_{\dot{x}^2=0}]^T$$

15

and

$$S(\dot{x}) = [1, \dot{x}^2, \ldots, (\dot{x}^2)^m]^T$$

The remainder ϵ_t is given by :

$$\epsilon_t = \frac{1}{(m+1)!} \frac{\vartheta^{(m+1)}\alpha}{\vartheta(\dot{x}^2)^{(m+1)}}|_{\dot{x}^2=0}(\dot{x}^2)^{m+1}$$

In theory, the remainder decreases as m increases. By numerical calculation, it is found that m needs to be very large in order to obtain an acceptable approximation accuracy, and that because c_2 is in the order of 10^{-4} in the denominators of the terms.

An another approximation method is the non-linear functional approximation. In this method under the assumption that $c_2 = \dot{x}_s^2$ is known exactly, (2.27) can be expanded around the nominal value c_{1n} using Taylor serias as:

$$\alpha(\dot{x}) = \sum_{k=0}^{m} \frac{\vartheta^m \alpha}{\vartheta c_1^m}|_{c_1=c_{1n}}(c_1 - c_{1n})^m + \epsilon_t = W^T S(\dot{x}) + \epsilon_t \qquad (2.29)$$

where :

$$W = [c_0, \ldots, (-1)^m c_0(c_1 - c_{1n})^m]^T$$

$$S(\dot{x}) = \Big[\frac{1}{1 + c_{1n}e^{-\dot{x}^2/c_2}}, \ldots, \frac{e^{-m\dot{x}^2/c_2}}{(1 + c_{1n}e^{-\dot{x}^2/c_2})^{m+1}}\Big] \qquad (2.30)$$

c_{1n} is the nominal value of c_1 and the remainder ϵ_t is given by, $\forall (0 \leq \xi \leq 1)$

$$\epsilon_t = \frac{(-1)^{(m+1)}c_0 e^{-(m+1)\dot{x}^2/c_2}}{(1 + \xi e^{-\dot{x}^2/c_2})^{m+1}}(c_1 - c_{1n})^{(m+1)} \qquad (2.31)$$

It can be seen that $|\epsilon_t| \leq c_0|c_1 - c_{1n}|^{(m+1)}, \forall c_1 \in (0,1)$. Because $|c_1 - c_{1n} < 1$, the upper bound $c_0|c_1 - c_{1n}|^{(m+1)}$ decreases as m increases. The result is global because the upper bound of the approximation error is independent of the operating range of $\dot{x}$, which can be very large. From (2.30) the primitives used to construct the basis function S are $e^{-\dot{x}^2/c_2}$ and $\frac{1}{1+c_{1n}e^{-\dot{x}^2/c_2}}$.

Each element s_i of S can be written as:

$$\Big[e^{-\dot{x}^2/c_2}\Big]^i \Big[\frac{1}{1 + c_{1n}e^{-\dot{x}^2/c_2}}\Big]^j = \frac{e^{-i\dot{x}^2/c_2}}{(1 + c_{1n}e^{-\dot{x}^2/c_2})^j}$$

16

with $i + j \geq 0, i, j \geq 0$. In general, the smaller the $i + j$, the more important this term is for the reconstruction. The higher the $i + j$ is, the better the approximation accuracy. If both c_1 and c_2 are unknown, then the non-linear function can be expanded with respect to both of them around their nominal values (c_{1n}, c_{2n}), subsequently the basis functions vector $S(\dot{x})$ and the corresponding unknown parameters W can be found. The remainder can be quantified similarly. By some calculations the primitives can be found for this case consist of: $\dot{x}^2, e^{-\dot{x}^2/c_{2n}}$ and $\frac{1}{1+c_{1n}e^{-\dot{x}^2/c_{2n}}}$. Each subelement s_i of S should be $\frac{\dot{x}^{2i}e^{-j\dot{x}^2/c_{2n}}}{(1+c_{1n}e^{-\dot{x}^2/c_{2n}})^k}$ with $i + j + k \geq 0, i, j, k \geq 0$. As before, the smaller $i + j$ is the more important this term is for the reconstruction. The higher the $i + j$ is the better the approximation accuracy.

If no knowledge is available a Neural Network as mentioned earlier can be used to generate I/O maps using the property that a multi-layer NN can approximate any function, under mild assumptions with any desired accuracy. It has been proven that any continuous functions not necessarily infinitely smooth, can be uniformly approximated by linear combinations of Gaussians. The Gaussian RBF neural network is a particular network architecture which uses l-numbers of Gaussian functions of the form:

$$s_i(x, \dot{x}) = \exp\left[-\frac{(x - \mu_{1i})^2 + (x - \mu_{2i})^2}{\sigma^2}\right]$$

where $[x, \dot{x}]^T \in \Re^2$ is the input variable, $\sigma^2 \in \Re$ is the variance and $[\mu_{1i}, \ \mu_{2i}]^T \in \Re$ is the center vector. The s_i's are the elements of the basic function vector $S(x, \dot{x})$ of the approximation:

$$\alpha_\alpha(w, x, \dot{x}) = W^T S(x, \dot{x})$$

The shortcoming of [9], [10] model, called as LuGre model, lies in the inadequacy of the hysteresis part since it does not account for non-local memory and it cannot accomodate arbitrary displacement-force transition curves.

In [22] noted that LuGre model whereas allows good description of the constant velocity behavior and offers a smooth transition at velocity reversal, the modeling capabilities in presliding regime are restricted as follows:

$\Rightarrow$ The model is too dissipative in presliding

$\Rightarrow$ The shape of the transition curve is fixed by the model and therefore cannot be adapted to actually measured values.

For the latter, beside the parameter σ_0 which models the initial stiffness at velocity reversal, no parameters are left for the shaping of the transition curves which will always have the same form and therefore is inadequate for fitting transition curves of arbitrary forms. In [22] is also presented an improved friction model which the friction force F is modelled by a set of two equations where as in the case of the LuGre model, depend on a state variable z representing the average deformation of the asperities of the contacting surfaces. The first equation, the friction force equation, is:

$$F = F_h(z) + \sigma_1 \frac{dz}{dt} + \sigma_2 v \tag{2.32}$$

where σ_1 is a micro-viscous damping coefficient, σ_2 is the viscous damping coefficient and v is the velocity of the moving object. $F_h(z)$ is the hysteresis friction force that is the part of friction force exhibiting hysteretic behavior. It is a static hysteresis nonlinearity with non-local memory. This hysteresis function is consisting of transition curves (curves between two reversal points or extrema). Each velocity reversal initiates a new transition curve, adds a new extremum to the hysteresis memory and resets the state variable z to zero. The transition curve which is active at a certain time is represented by $F_d(z)$ and the value of $F_h(z)$ at the beginning of a transition curve is represented by F_b. Then we have:

$$F_h(z) = F_b + F_d(z) \tag{2.33}$$

F_d as they mention, is a point-symmetrical strictly increasing function of z.

The second equation, the nonlinear equation, is based on the current hysteresis transition curve $F_d(z)$ and the current velocity, that is:

$$\frac{dz}{dt} = v\left(1 - sign(\frac{F_d(z)}{S(v) - F_b}) * |\frac{F_d(z)}{S(v) - F_b}|^n\right) \qquad (2.34)$$

where $S(v)$ is the constant velocity behavior in sliding which is the same as in (2.17) without accounting for viscous friction term, that is $\sigma_2 v$. The parameter n, allows to modify the influence of $F_d(z)/(S(v) - F_b)$ on the difference between dz/dt and v such that the model behavior correspond better to friction measurements in the transition from presliding to sliding. The constant velocity behavior (sliding) where is described by this model is exactly the same as in the LuGre model. Namely, the friction force is given by:

$$F = F_b + F_d(z) + \sigma_2 v = S(v) + \sigma_2 v \qquad (2.35)$$

where is in the same form as in (2.17). The difference over the LuGre model consists in the zero velocity behavior (presliding) where a hysteresis model with non-local memory is included. Consequently from (2.32) and (2.34), for zero velocity we have :

$$\begin{aligned} F &= F_b + F_d(z) = F_h(z) \\ \frac{dz}{dt} &= 0 \end{aligned}$$

The hysteresis model relates the state variable z and the hysteresis force F_h. The implementation of the hysteresis model requires two memory stacks the one for the minima of F_h in ascending order (stack m) and one for the maxima of F_h (stack M). The stacks grow at velocity reversal and shrink when an internal hysteresis loop is closed. The stacks are reset when the system goes from presliding to sliding. The

19

value of F_b equals to the most recent element of stack M if the transition curve is descending and of stack m if the transition curve is ascending. The value of the state variable z is reset to zero at each velocity reversal and recalculated at the closing of an internal loop. In [22] is given an analytical description of the mechanisms which govern the hysteresis model, mechanisms that also exist in [23] and on the following hysteresis section of this report.

The above model can account accurately for experimentally obtained friction characteristics which are: Stribeck friction in sliding, hysteretic behavior in presliding, frictional lag, varying break-away and stick-slip behavior.

2.2   Backlash

Actuator and sensor nonlinearities are among the key factors limiting both static and dynamic performance of feedback control systems. Harmful effects of backlash in gears are well known. Backlash prevents accurate positioning and may lead to chattering and limit-cycle-type instabilities. This increases wear and tear of the gears, which, in turn, increases backlash. This phenomenon has haunted the constructors of control systems for more than 50 years: from the servomechanisms in the 1940s to the modern high precision robotic manipulators. Typically the concept of backlash is associated with gear trains and similar mechanical couplings. Sometimes backlash can be used to approximate description of the delays in drives with elastic cables or in long pipes.

2.2.1   Mathematical Models of Backlash

The most familiar and simple model perhaps is the one for backlash hysteresis (piecewise linear model) that the backlash characteristic $u(t) = B(v(t))$ described by two parallel lines connected via horizontal lines (see fig. 2.2). The methods that
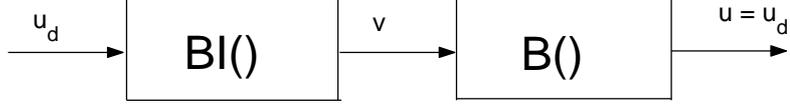
Figure 2.1: Inverting a backlash

are presented in [19] and [21], need to construct an inverse model to mitigate the effects of the backlash (see fig. 2.1).

Mathematically, this phenomenon is modelled as:

$$\dot{u}(t) = \begin{cases} m\dot{v}(t) & \text{if } \dot{v}(t) > 0 \text{ and } u(t) = m(v(t) - c_r) \\ & or \\ & \text{if } \dot{v}(t) < 0 \text{ and } u(t) = m(v(t) - c_l) \\ 0 & \text{otherwise} \end{cases} \qquad (2.36)$$

with input $v(t)$ and output $u(t)$ and c_l, c_r the left and right crossings respectively with $c_r > c_l$. The $v(t)$ and $\dot{v}(t)$ uniquely determine $u(t), \dot{u}(t)$ and the knowledge of $\dot{v}(t)$ is necessary to specify the signal motion of $B(\cdot)$ ($B(\cdot)$ the backlash characteristic) on whether a straight line or an inner segment (horizontal). A further insight into the nature of backlash can be gained from the waveforms of the output $u(t)$ when the input is $v(t)$ is a sine signal in fig. 2.2. For this illustration the backlash parameters are $m = 1$, $c_r = 0.5$, $c_l = -0.5$ and for three different initial conditions at a time, i.e. $u(0) = -0.5$, $u(0) = 0$ and $u(0) = 0.3$. It is noticeable that for the last two initials conditions we obtain initial "transients" and then $u(t)$ settle in its periodic steady state. For the initial condition $u(0) = -0.5$ the periodic steady state is reached without transients. The periodic steady state of $u(t)$ reveals the two fundamental features of backlash. First, it introduces a phase delay. Second, it causes a loss of information by chopping the peaks of $v(t)$.
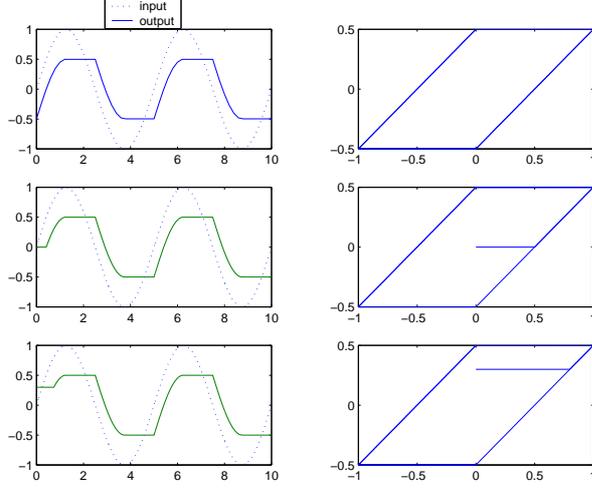
Figure 2.2: Backlash response to a sine input.

The inverse model of (2.36) as proposed is:

$$
\dot{v}(t) = \begin{cases}
\frac{1}{m}\dot{u}_d(t) & \text{if } \dot{u}_d(t) > 0 \text{ and } v(t) = \frac{1}{m}u_d(t) + c_r \\
& or \\
& \text{if } \dot{u}_d(t) < 0 \text{ and } v(t) = \frac{1}{m}u_d(t) + c_l \\
0 & \text{if } \dot{u}_d(t) = 0 \\
g(t,t) & \text{if } \dot{u}_d(t) > 0 \text{ and } v(t) = \frac{1}{m}u_d(t) + c_l \\
-g(t,t) & \text{if } \dot{u}_d(t) < 0 \text{ and } v(t) = \frac{1}{m}u_d(t) + c_r
\end{cases}
\tag{2.37}
$$

where $g(\tau, t) = \delta(\tau - t)(c_r - c_l)$ with $\delta(t)$ the Dirac δ-function. $u_d(t)$ is a given desired signal for $u(t)$ and a backlash inverse $BI(\cdot)$ defined by (2.37) is such that $u_d(t) = B(BI(u_d(t)))$.

In the above definition the inverse of a horizontal segment of the backlash characteristic is a vertical jump of a distance $(c_r - c_l)$. The following lemma that it is being proved in [21] is the following one:

Lemma 1

The characteristic $BI(\cdot)$ defined by (2.37) is the right inverse of the characteristic

22

(2.36) in the sense:

$B(BI(u_d(t_0))) = u_d(t_0) \Rightarrow B(BI(u_d(t))) = u_d(t), \forall t \geq t_0$ for any piecewise continuous $u_d(t)$ and any $t \geq t_0$.

It is also reported that an initialization of the backlash inverse $BI(\cdot)$ for $u(t) = u_d(t)$ is possible at any given time t_0. When the parameters m, c_r, c_l are unknown instead of these, are used their estimates $\hat{m}(t), \hat{c}_r(t), \hat{c}_l(t)$ to design an adaptive feedback backlash inverse. As before, $u_d(t), \dot{u}_d(t)$ uniquely determine $v(t), \dot{v}(t)$ and the knowledge of $\dot{u}_d(t)$ is necessary to specify the signal motion of the backlash inverse.

In [13] is defined a dynamic hysteresis model to approximate the backlash hysteresis, an approximation which is called backlash-like hysteresis. As in [21] a backlash hysteresis non-linearity can be described by:

$$w(t) = P(v(t)) = \begin{cases} c(v(t) - B) & \text{if } \dot{v}(t) > 0 \text{ and } w(t) = c(v(t) - B) \\ c(v(t) + B) & \text{if } \dot{v}(t) < 0 \text{ and } w(t) = c(v(t) + B) \\ w(t_-) & \text{otherwise} \end{cases} \qquad (2.38)$$

where $c > 0$ is the slope of the lines and $B > 0$ is the backlash distance. The continuous-time dynamic model which describes a class of backlash-like hysteresis is given by the following equation:

$$\frac{dw}{dt} = \alpha \left| \frac{dv}{dt} \right| (cv - w) + B_1 \frac{dv}{dt} \qquad (2.39)$$

where α and B_1 are constants satisfying $c > B_1$. The equation (2.39) can be solved explicitly for v piecewise monotone:

$$w(t) = cv(t) + d(v) \qquad (2.40)$$

with :

$$d(v) = [w_0 - cv_0]e^{-\alpha(v-v_0)sgn\dot{v}} + e^{-\alpha v sgn\dot{v}} \int_{v_0}^{v} [B_1 - c]e^{\alpha\zeta(sgn\dot{v})}d\zeta \qquad (2.41)$$

23

for $\dot{v}$ constant, $w(v_0) = w_0$. For $d(v)$, it can be shown that if $w(v; w_0, v_0)$ is the solution of (2.39) with initial values (v_0, w_0) then if $\dot{v} > 0 (\dot{v} < 0)$ and $v \to \infty(-\infty)$, one has:

$$\lim_{v \to \infty} d(v) = \lim_{v \to \infty} [w(v; v_0, w_0) - f(v)] = -\frac{c - B_1}{\alpha} \qquad (2.42)$$

$$\lim_{v \to -\infty} d(v) = \lim_{v \to -\infty} [w(v; v_0, w_0) - f(v)] = \frac{c - B_1}{\alpha} \qquad (2.43)$$

The above convergence is exponential at the rate of α. Solution (2.40) and properties (2.42),(2.43) show that $w(t)$ eventually satisfies the first and second conditions of (2.38). Moreover, setting $\dot{v} = 0$ results in $\dot{w} = 0$ which satisfies the last condition of (2.38). This implies that the dynamic equation (2.39) can be utilized to model a class of backlash-like hysteresis and is an approximation of backlash hysteresis (2.38). Equations (2.42) and (2.43) indeed show that w switches exponentially from the line $cv(t) - \frac{c - B_1}{\alpha}$ to $cv(t) + \frac{c - B_1}{\alpha}$ to generate backlash-like hysteresis curve. The solutions of (2.39) can be obtained by numerical integration with v as the independent variable. The parameter α determines the rate for $w(t)$ to switch between $-\frac{c - B_1}{\alpha}$ and $\frac{c - B_1}{\alpha}$. The larger the parameter α is, the faster the transition in $w(t)$ is going to be. However the backlash distance is determined by $\frac{c - B_1}{\alpha}$ and the parameter must satisfy $c > B_1$. Consequently a compromise should be made for choosing a suitable parameter set $\{\alpha, c, B_1\}$ to model the required shape of backlash-like hysteresis. If the values of backlash slope and distance are not exactly known, then adaptations will be used to estimate them.

The useful outcome of [13] is that backlash-like hysteresis is modelled by a dynamic equation without the need to construct a backlash hysteresis inverse.

It is also reported that in the presence of actuator dynamics prior to the backlash, the adaptive backlash inverse control problem is more difficult because it requires that backlash be inverted through a dynamic block and this problem is currently under investigation.

2.3 Hysteresis

As it is characterized by the literature, backlash is the simplest form of hysteresis. Hysteresis phenomena are even more numerous and diverse than those modeled by backlash characteristics. Generally include nondifferentiable nonlinearities and usually unknown. While ferromagnetic hysteresis is the best known type of hysteresis, similar characteristics are common in plastic, piezoelectric and other materials. However it is in general unrealistic to expect that a single hysteresis model can serve a vast variety of applications. In the sequel of this section various models of hysteresis are presented. In fig. 2.3 is given a typical graph of hysteresis phenomenon.
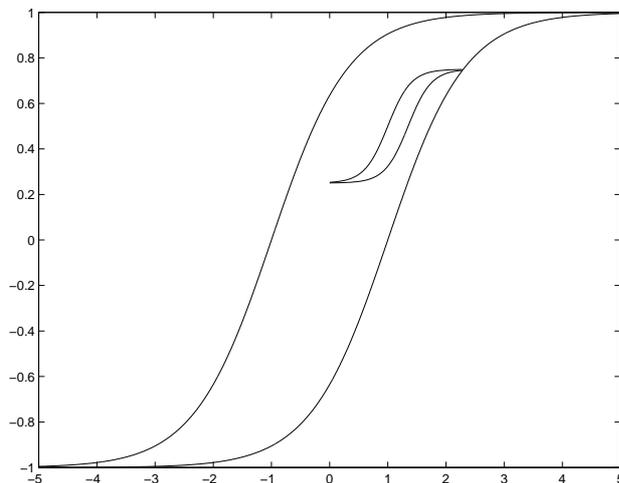
Figure 2.3: A typical Hysteresis diagram with a major and minor loop

2.3.1 Mathematical Models of Hysteresis

In [20] a simplified hysteresis model is used that captures most of the hysteresis characteristics and is useful for parameter adaptive control. It has been showed that the proposed model has a parameterizable right inverse which cancels the effect of

the hysteresis when cascaded with the hysteresis. Its main hysteresis and two minor loops are shown in fig. 2.4. It can be tuned by as many as eight parameters:
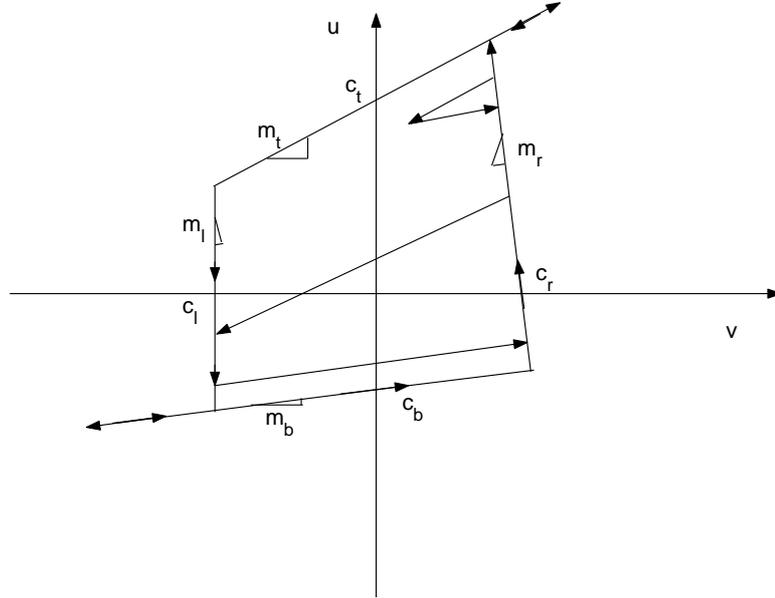
Figure 2.4: Hysteresis model

four slopes m_l, m_r, m_b, m_t and four crosssing parameters c_l, c_r, c_b, c_t, where the subscripts indicate "left", "right", "bottom" and "top" respectively. The difference between the slopes m_b and m_t allows for the appearance of local loops.

Defining as:

$$v_1 \triangleq \frac{c_t + m_l c_l}{m_l - m_t} \qquad v_2 \triangleq \frac{c_b + m_r c_r}{m_r - m_b}$$
$$v_3 \triangleq \frac{c_t + m_r c_r}{m_r - m_t} \qquad v_4 \triangleq \frac{c_b + m_l c_l}{m_l - m_b}$$

where v_1, v_2, v_3, v_4 are the values of $v(t)$ at the upper-left, lower-right, upper-right and lower-left corners of the quadrilateral. Then the hysteresis $u(t) = H(v(t))$ representing the motion of u(t) and v(t) is fully described by:

26

$$
\dot{u}(t) = \begin{cases}
m_t \dot{v}(t) & \text{if } v(t) \geq v_3 \ \mathcal{AND} \ u(t) = m_t v(t) + c_t, \ \mathcal{OR} \\
& \text{if } v_4 < v(t) < v_3, \dot{v}(t) < 0, \\
& u(t) = m_t v(t) + c_d, \ u(t) \neq m_l(v(t) - c_l) \ \mathcal{AND} \ u(t) \neq m_b v(t) + c_b, \ \mathcal{OR} \\
& \text{if } m_t < m_b, \ v_4 < v(t) < v_3, \\
& u(t) = m_b v(t) + c_b \ \mathcal{AND} \ \dot{v}(t) < 0, \ \mathcal{OR} \\
& \text{if } m_t < m_b, \ v_4 < v(t) < v_3, \\
& u(t) = m_t v(t) + c_t \ \mathcal{AND} \ \dot{v}(t) > 0 \\[1em]
m_b \dot{v}(t) & \text{if } v(t) \leq v_4 \ \mathcal{AND} \ u(t) = m_b v(t) + c_b, \ \mathcal{OR} \\
& \text{if } v_4 < v(t) < v_3, \dot{v}(t) > 0, \\
& u(t) = m_b v(t) + c_u, \ u(t) \neq m_r(v(t) - c_r) \ \mathcal{AND} \ u(t) \neq m_t v(t) + c_t, \ \mathcal{OR} \\
& \text{if } m_t > m_b, \ v_4 < v(t) < v_3, \\
& u(t) = m_t v(t) + c_t \ \mathcal{AND} \ \dot{v}(t) > 0, \ \mathcal{OR} \\
& \text{if } m_t > m_b, \ v_4 < v(t) < v_3, \\
& u(t) = m_b v(t) + c_b \ \mathcal{AND} \ \dot{v}(t) < 0 \\[1em]
m_r \dot{v}(t) & \text{if } v_4 < v(t) < v_3, \ \dot{v}(t) > 0 \ \mathcal{AND} \\
& u(t) = m_r(v(t) - c_r) \\[1em]
m_l \dot{v}(t) & \text{if } v_4 < v(t) < v_3, \ \dot{v}(t) < 0 \ \mathcal{AND} \\
& u(t) = m_l(v(t) - c_l) \\[1em]
0 & \text{if } \dot{v}(t) = 0
\end{cases}
$$

$$(2.44)$$

These expressions indicate that is hysteresis is a complex nonlinear dynamic system defined by piecewise linear relationships between the input $v(t)$, output $u(t)$, and their time derivatives.

With $u_d(t)$ be a control signal to be designed, the inverse of the hysteresis, name it $HI(\cdot)$, it is given by the motion of $u_d(t)$ and $v(t)$ and mathematically is described

27

as:

$$
\dot{v}(t) = \begin{cases}
\frac{1}{m_t}\dot{u}_d(t) & \text{if } u_d(t) \geq u_3, \; \mathcal{OR} \\
& \text{if } u_4 < u_d(t) < u_3, \; \dot{u}_d(t) < 0, \\
& v(t) \neq \frac{1}{m_l}u_d(t) + c_l \; \mathcal{AND} \; v(t) \neq \frac{1}{m_b}(u_d(t) - c_b), \; \mathcal{OR} \\
& \text{if } m_t < m_b, \; u_4 < u_d(t) < u_3, \\
& v(t) = \frac{1}{m_b}(u_d(t) - c_b) \; \mathcal{AND} \; \dot{u}_d(t) < 0, \; \mathcal{OR} \\
& \text{if } m_t < m_b, \; u_4 < u_d(t) < u_3, \\
& v(t) = \frac{1}{m_t}(u_d(t) - c_t) \; \mathcal{AND} \; \dot{u}_d(t) > 0 \\[1em]
\frac{1}{m_b}\dot{u}_d(t) & \text{if } u_d(t) \leq u_4, \; \mathcal{OR} \\
& \text{if } u_4 < u_d(t) < u_3, \dot{u}_d(t) > 0, \\
& v(t) \neq \frac{1}{m_r}u_d(t) + c_r \; \mathcal{AND} \; v(t) \neq \frac{1}{m_t}(u_d(t) - c_t), \; \mathcal{OR} \\
& \text{if } m_t > m_b, \; u_4 < u_d(t) < u_3, \\
& u(t) = \frac{1}{m_t}(u_d(t) - c_t) \; \mathcal{AND} \; \dot{u}_d(t) > 0, \; \mathcal{OR} \\
& \text{if } m_t > m_b, \; u_4 < u_d(t) < u_3, \\
& u(t) = \frac{1}{m_b}(u_d(t) - c_b) \; \mathcal{AND} \; \dot{u}_d(t) < 0 \\[1em]
\frac{1}{m_r}\dot{u}_d(t) & \text{if } u_4 < u_d(t) < u_3, \; \dot{u}_d(t) > 0 \; \mathcal{AND} \\
& v(t) = \frac{1}{m_r}u_d(t) + c_r \\[1em]
\frac{1}{m_l}\dot{u}_d(t) & \text{if } u_4 < u_d(t) < u_3, \; \dot{u}_d(t) < 0 \; \mathcal{AND} \\
& v(t) = \frac{1}{m_l}u_d(t) + c_l) \\[1em]
0 & \text{if } \dot{u}_d(t) = 0
\end{cases}
\tag{2.45}
$$

with:

$$
u_1 \triangleq \frac{m_l(m_t c_l + c_t)}{m_l - m_t} \qquad u_2 \triangleq \frac{m_r(m_b c_r + c_b)}{m_r - m_b}
$$
$$
u_3 \triangleq \frac{m_r(m_t c_r + c_t)}{m_r - m_t} \qquad u_4 \triangleq \frac{m_l(m_b c_l + c_b)}{m_l - m_b}
$$

The proposed hysteresis inverse has the following property:

Proposition 1

The characteristic $HI(\cdot)$ given by (2.45) is the right inverse of the characteristic (2.44) in the sense:

$H(HI(u_d(t_0))) = u_d(t_0) \Rightarrow H(HI(u_d(t))) = u_d(t), \forall t \geq t_0$ for any piecewise continuous $u_d(t)$ and any $t \geq t_0$.

However, as $u_d(t)$ is the design signal of our choice, an initialization of the hysteresis inverse by an appropriate choice of $u_d(t_0)$ should always make $v(t)$ and $u(t)$ leave the inside loop at t_0 so that $u(t_0) = u_d(t_0)$ and then from Proposition 1, $u(t) = u_d(t)$ for any $t \geq t_0$. When the parameters m_t, c_t, m_b, c_b, m_r, c_r, m_l, c_l are unknown instead of these, are used their estimates $\hat{m}_t$, $\hat{c}_t$, $\hat{m}_b$, $\hat{c}_b$, $\hat{m}_r$, $\hat{c}_r$, $\hat{m}_l$, $\hat{c}_l$ to design an adaptive feedback hysteresis inverse.

In [11] is defined a mathematical model for hysteresis loop. This model is a first order nonlinear differential equation and is capable, as they allegate, of simulating exactly a given hysteresis loop and furthermore the model exhibits many of the properties that are, in fact, observed in practice. The mathematical model is the following one :

$$\frac{dy}{dt} = h(y)g \circ [x(t) - f(y)] \tag{2.46}$$

where f, g, h are real-valued functions defined on the real line $\Re$ and '$\circ$' denotes the composition operation. The set of functions having K continuous derivatives on $\Re$ are denoted by $C^k(\Re)$. The three functions are assumed to satisfy the conditions:

(i) $g, f, h \in C^1(\Re)$

(ii) $g' > 0, f' > 0$ on $\Re$

(iii) $f, g : \Re \to \Re$

(iv) $0 < a \leq h < b < \infty$ on $\Re$

where a, b are finite positive constants. The function g is referred to as the dissipation function, the function f as the restoring function and h is called the weighting function. This model, given by a non-linear differential equation, represents a dynamic

29

process. The model exhibits many of the observed properties, such as widening effects with increasing frequency and minor hysteresis loops when a periodic signal is superimposed upon a constant signal.

They are also reported some important properties that govern the model (2.46) where the proofs here have been deleted for brevity. Primarily let denote I the interval on the real line $[0, \infty)$. A solution to (2.46) with initial condition $y(t_0) = y_0$ where $t_0 \in I$ is a differentiable function ϕ defined on the interval I such that:

$$\phi(t) = h(\phi(t))g \circ [x(t) - f(\phi(t))] \text{for } t \geq 0 \text{ and } \phi(t_0) = y_0 \qquad (2.47)$$

The six properties are the following:

- **Property 1** If $x(t)$ is bounded and continuous on I, the for all $y_0 \in \Re$ and all $t_0 \geq 0$

    (a) there exists a solution $\phi(t)$ satisfying (2.46) for all $t \geq 0$

    (b) $\phi(t)$ is uniformly bounded

    (c) $\phi(t)$ is unique

- **Property 2** If $x(t)$ is bounded, continuous on I and periodic of fundamental period T, then there exists a unique periodic solution $p(t)$ to equation (2.46) of the same fundamental period T.

- **Property 3** With $x(t)$ satisfying the same conditions as in property 2 any solution $\phi(t)$ to equation (2.46) with arbitrary initial conditions will approach asymptotically the unique periodic solution $p(t)$

- **Property 4** If $x(t)$ is bounded, continuous on I and periodic of fundamental period T and if in addition $x(t)$ has only one maximum and one minimum per cycle, with no points of inflection, then the unique periodic solution $p(t)$ of equation (2.46) also has only one maximum and one minimum per cycle.

- **Property 5** If $x(t)$ satisfies the condition of property 4 and let $p(t)$ be the unique periodic solution to equation (2.46), then a parameterized curve defined by :

$$\Gamma = \{(x, y) : x = x(t), y = p(t), \ 0 \leq t \leq T\}$$

  is a simple closed curve.

- **Property 6** if $y_2(t) = y_1(\alpha t)$ with $0 \leq \alpha \leq 0$ and $y_1(t) = y_1(t + T_1)$ and the two associated simple closed curves defined by :

$$\Gamma_1 = \{(x, y) : x = x(t), y = y_1(t), \ 0 \leq t < T_1\}$$

$$\Gamma_2 = \{(x, y) : x = x(t), y = y_2(t), \ 0 \leq t < T_1/\alpha\}$$

  Then the area enclosed by the curve Γ_1 is greater than the area enclosed by Γ_2.

The one that remains is to present a procedure for constructing the dissipation function g, the restoring function f and the weighting function h. In order to identify these functions a pair of waveforms $\{x(t), y(t)\}$ must be measured. Because the system characterized as nonlinear there is no apparent advantage in using one set of measurements over another, unlike the linear case. The authors report that one procedure is to select a pair of waveforms $\{x(t), y(t)\}$ where $y(t)$ is a cosine function. Namely to apply a cosine waveform $y(t)$ and measuring the corresponding waveform $x(t)$. With $y(t)$ constrained to be cosine waveform it follows that $\dot{y}(t)$ is known and (2.46) is reduced to an algebraic relationship.

For a given closed hysteresis loop where $y(t)$ is a cosine function, g can be chosen as an arbitrary odd function providing it satisfies the necessary conditions $(i) - (iv)$ where are mentioned above, and construct f, h so that equation (2.46) can be used

to represent exactly a given hysteresis loop. Since both $y(t)$ and $\dot{y}(t)$ are known and since g^{-1} exists we can rewrite equation (2.46) as :

$$x(t) = g^{-1}\left(\frac{\dot{y}(t)}{h(y(t))}\right) + f(y(t)) \tag{2.48}$$

With $y(t)$ a cosine function of period of T for each y in the range of $y(t)$ (except for the extremal points) there exists two values of $t \in [0, T]$ say t_1, t_2 where $y(t_1) = y(t_2) = y$. It is obvious that $\dot{y}(t_1) = -\dot{y}(t_2)$. Thus with g odd

$$g^{-1}\left(\frac{\dot{y}(t_1)}{h(y(t_1))}\right) = -g^{-1}\left(\frac{\dot{y}(t_2)}{h(y(t_2))}\right) = d \tag{2.49}$$

and since both $(x(t_1), y(t_1))$ and $(x(t_2), y(t_2))$ represent points on the hysteresis loop with same ordinate $x(t_1) - x(t_2) = 2d$. Moreover, in view of equations (2.48), (2.49) the midpoint of the two points on the hysteresis loop corresponding to t_1 and t_2 satisfies the equation:

$$y(t_1) = y(t_2) = f^{-1}\left(\frac{x(t_1) + x(t_2)}{2}\right) \tag{2.50}$$

where $\left(\frac{x(t_1)+x(t_2)}{2}\right)$ is on the midpoint of the hysteresis loop taken along the abscissa direction. Thus the locus of midpoints of the hysteresis loop determines the f^{-1} function. Since g (and hence g^{-1}) is known and since the value of d can be measured directly from the hysteresis loop, $h(y(t_1))$ can be obtained from equation (2.49). Continuing for all values of $t \in [0, T)$ the function h can be determined. The points from the above description are given in fig. 2.5.

The functions f, h constructed by this procedure are unique. The arbitrary choice of g is an advantage which permits us to represent more closely a family of hysteresis loops. Hence if we are attempting to model a family of hysteresis loops rather than a single loop, appropriate optimization techniques may be used to determined the functions f, g, h.

The main shortcoming from the above discussed model that is not predicted is its
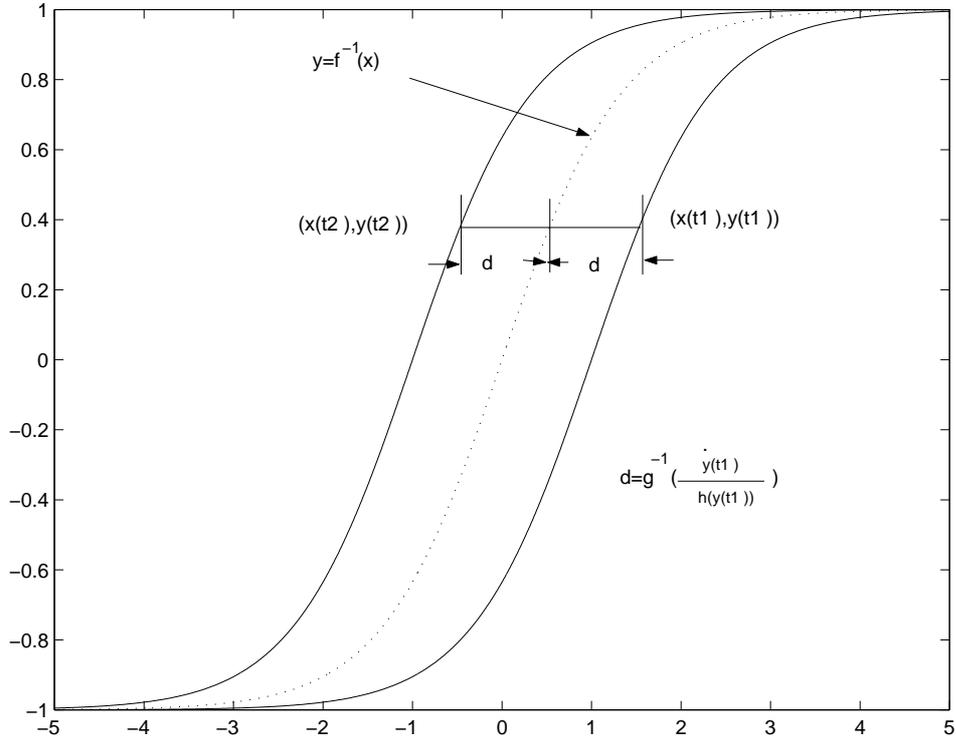
Figure 2.5: Construction procedure for obtaining f,g and h

behavior when x is constant (called as dc behavior) where y assume more than one distinct state. To overcome this drawback an improved model is presented in [12] where the dc hysteresis is included while it preserves the good characteristics of the model (2.46) namely unusual versatility, loop widening with increasing frequency and the lack of any required special handling once the model is placed in the system. It also exhibits loop narrowing with increasing frequency, a phenomenon in the $i - v$ curves of fluorescent lamps, such as reduction reduction in loop widening to insignificant amounts of beyond an upper threshold frequency while including strong widening at intermediate frequencies and a dc loop (phenomenon which is presented in iron-core materials). This hysteresis model is the following :

$$\frac{dy}{dt} = w \circ \left(\frac{dx}{dt}\right) h \circ (y(t)) g \circ (x(t) - f(y(t))) \qquad (2.51)$$

where as before '$\circ$' denotes functional composition. The new characteristic to this model is the function w where it is also assumed that belongs to class $C^1(\Re)$ and $w \circ (dx/dt) \geq 0$. Also is assumed that $w \circ (dx/dt) = 0 \Leftrightarrow dx/dt = 0$. Moreover, the model which is given by (2.51) has the similar properties of that of the model which is given by (2.46). In fig. 2.6 several possible w functions are shown along with brief descriptions of their effects upon the friction behavior of (2.51).

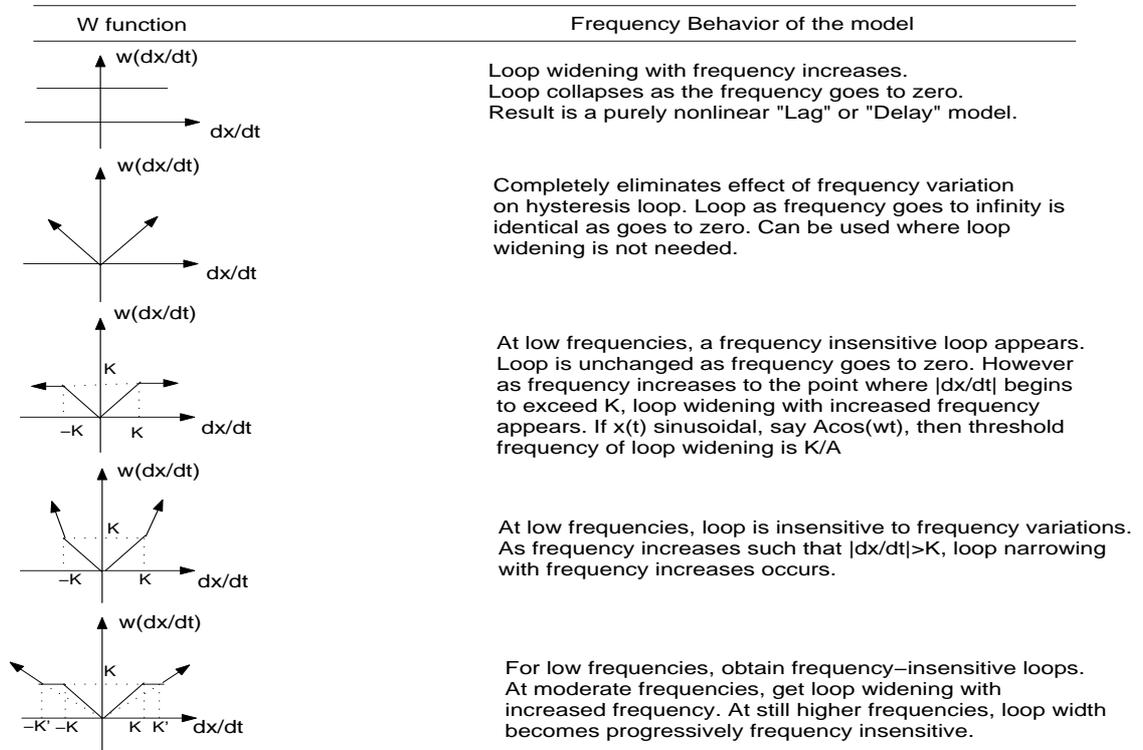| W function	Frequency Behavior of the model
w(dx/dt) ... dx/dt	Loop widening with frequency increases. Loop collapses as the frequency goes to zero. Result is a purely nonlinear "Lag" or "Delay" model.
w(dx/dt) ... dx/dt	Completely eliminates effect of frequency variation on hysteresis loop. Loop as frequency goes to infinity is identical as goes to zero. Can be used where loop widening is not needed.
w(dx/dt) ... K ... −K K dx/dt	At low frequencies, a frequency insensitive loop appears. Loop is unchanged as frequency goes to zero. However as frequency increases to the point where |dx/dt| begins to exceed K, loop widening with increased frequency appears. If x(t) sinusoidal, say Acos(wt), then threshold frequency of loop widening is K/A
w(dx/dt) ... K ... −K K dx/dt	At low frequencies, loop is insensitive to frequency variations. As frequency increases such that |dx/dt|>K, loop narrowing with frequency increases occurs.
w(dx/dt) ... K ... −K' −K K K' dx/dt	For low frequencies, obtain frequency–insensitive loops. At moderate frequencies, get loop widening with increased frequency. At still higher frequencies, loop width becomes progressively frequency insensitive.

Figure 2.6: Possible w functions and their effects on frequency behavior

We should note that these piecewise linear w functions may be replaced with smooth versions if necessary. The basic drawback of (2.51) is, of course the need for introduction of parasitics under arbitrary excitation. At very low frequencies, this can

drastically increase computer solution time.

An another approach which is prevalent in hysteresis modelling is the Preisach-type models of hysteresis. All the models have a common generic feature; they are constructed as superpositions of simplest hysteresis nonlinearities-rectangular loops. In the following discussion, according to [23] here are reported various generalizations and extensions of the classical Preisach model, giving the necessary and sufficient conditions for the representation of actual hysteresis nonlinearities by various Preisach-type models, the solution of identification problems for these models and numerical implementation.

Starting with the definition of scalar hysteresis we consider a transducer (see fig. 2.7) which is called a hysteresis transducer (HT) if its I/O relationship is a multibranch

$u(t)$ HT $f(t)$

Figure 2.7: Hysteresis Transducer

nonlinearity for which branch-to-branch transitions occur after input extrema. The term *static* hysteresis nonlinearity means that the branches are determined only by the past extremum values of input while the speed of input variations between extremum points has no influence on branching. It is worthwhile to keep in mind that, for very fast input variations, time effects become important and the given definition of static hysteresis fails. It is also important to mention that the given definition of hysteresis emphasizes the fact that branching constitutes the essence of hysteresis, while the formation of hysteresis loops (looping) is a particular case of branching. Indeed, looping occurs when the input varies back and forth between two consecutive extremum values while branching takes place for arbitrary input variations. All static hysteresis fall into two general classifications :

(a) **Hysteresis nonlinearities with local memories.** The value of output $f(t_0)$ at some instant of time t_0 and the values of input $u(t)$ at all subsequent instants of time $t \geq t_0$ uniquely predetermine the value of output $f(t)$ for all $t > t_0$. In other words, the past exerts its influence upon the future through the current value of output.

(b) **Hysteresis nonlinearities with non-local memories.** In antithesis, for the *non-local* hysteresis nonlinearities the future values of output $f(t)$ $(t \geq t_0)$ depend not only on the current value of output $f(t_0)$ but on past extremum values of input as well.

It is clear from the above that all hysteresis nonlinearities with local memories have the following common feature:

Every reachable point in the $f - u$ diagram corresponds to a uniquely defined state. This state predetermines the behavior of HT in exactly one way for increasing $u(t)$ and in exactly one way for decreasing $u(t)$. Namely, at any point in the $f - u$ diagram there are only one or two curves that may represent the future behavior of HT with local memory.

From the other hand in the case of non-local memories, at any reachable point in the $f - u$ diagram there is an infinity of curves that may represent the future behavior of the transducer. Each of these curves depends on a particular past history, namely, on a particular sequence of past extremum values of input.

To describe the mathematical Preisach model we consider an infinite set of simplest hysteresis operators $\hat{\gamma}_{\alpha\beta}$. Each of these operators can be represented by a rectangular loop on the I/O diagram (see fig. 2.8). Numbers α and β correspond to "up" and "down" switching values of input respectively. It is assumed that $\alpha \geq \beta$. It is apparent that these operators $\hat{\gamma}_{\alpha\beta}$ represent hysteresis nonlinearities with local memories. Along with the set of operators $\hat{\gamma}_{\alpha\beta}$ consider an arbitrary weight function $\mu(\alpha, \beta)$ referred to as the Preisach function. Then the Preisach model can be written
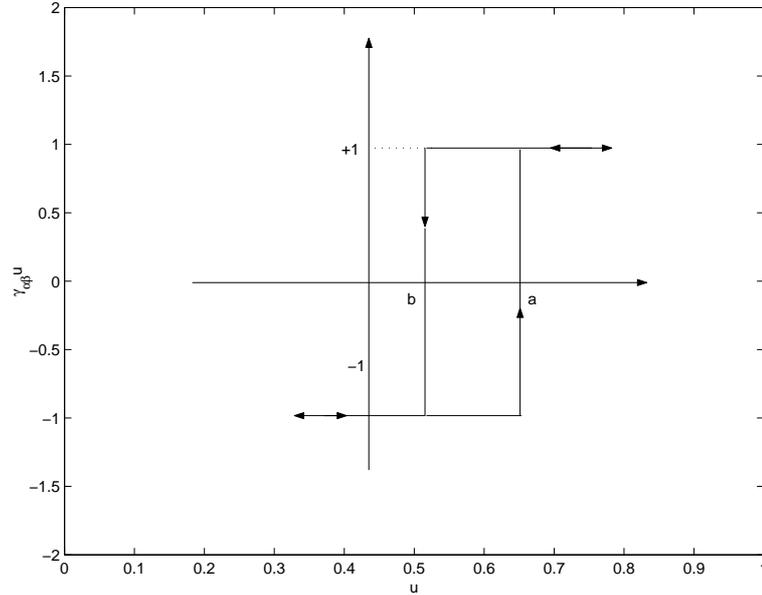
Figure 2.8: A simple hysteresis operator $\hat{\gamma}_{\alpha\beta}$

as:

$$f(t) = \hat{\Gamma}u(t) = \int\int_{\alpha \geq \beta} \mu(\alpha, \beta)\hat{\gamma}_{\alpha\beta}u(t)d\alpha d\beta \qquad (2.52)$$

Although the Preisach hysteresis nonlinearity (2.52) is constructed as a superposition of elementary hysteresis nonlinearities $\hat{\gamma}_{\alpha\beta}$ with local memories it usually has non-local memory. Introducing now, the model's geometric interpretation we can see that there is one to one correspondence between operators $\hat{\gamma}_{\alpha\beta}$ and points (α, β) of the half-plane (see fig. 2.9) $\alpha \geq \beta$. Consequently each point of the half-plane $\alpha \geq \beta$ can be identified with only one particular $\hat{\gamma}$-operator whose "up" and "down" switching values are respectively equal to α and β coordinates of the point. It is also assumed that the Preisach function equals to zero outside of the triangle T (see fig. 2.9). It is worthwhile to give an example explaining the usefulness of the Preisach phase plane. Assuming that input $u(t)$ at some instant of time t_0 has the value which is less than β_0. Then the outputs of all $\hat{\gamma}$-operators which correspond to

37
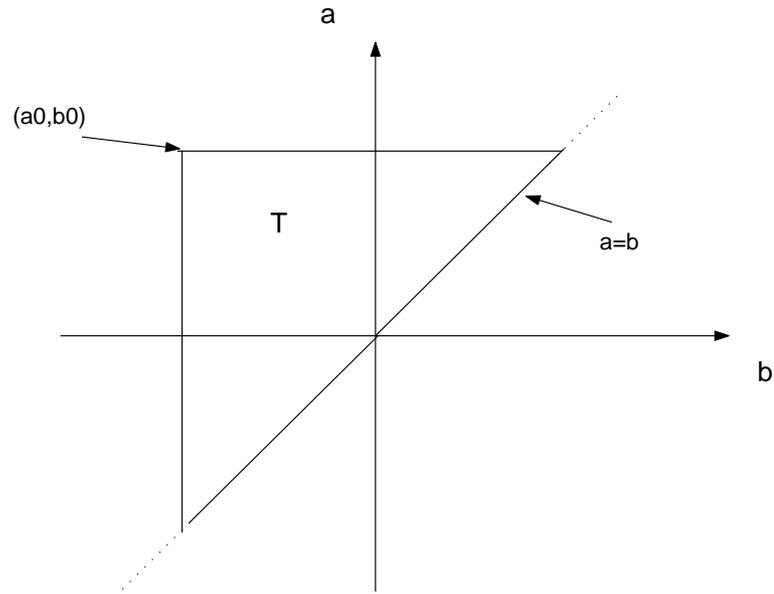
Figure 2.9: The Preisach phase plane

the points of the triangle T are equal to -1 (negative saturation). Then we assume that the input increases monotonically until it reaches at time t_1 some maximum value u_1. Then decreases monotonically until it reaches some minimum value u_2 at time instant t_2. Geometrically the triangle T it winnows into the two sets: $S^+(t)$ consisting of points (α, β) for which $\hat{\gamma}$-operators are in the "up"-position and $S^-(t)$ consisting of points (α, β) for which $\hat{\gamma}$-operators are in the "down"-position (see fig. 2.10). Lets name $L(t)$, as the interface between $S^+(t)$ and $S^-(t)$. The above discussion reveals the mechanism of memory formation in the Preisach model. The memory is formed as result of two different rules for the modification of the interface $L(t)$. For a monotonically increasing input we have a horizontal final link of $L(t)$ moving upward, while for a monotonically decreasing input we have a vertical final link of $L(t)$ moving from right to the left. These two different rules result in the formation of the staircase interface $L(t)$ whose vertices have coordinates equal to past input extrema.
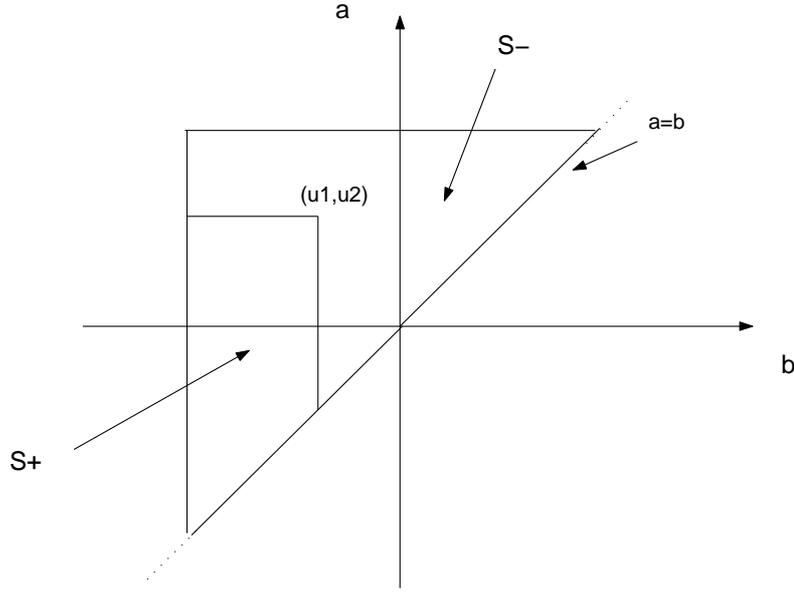
Figure 2.10: The formation of S^+ and S^-

From (2.52) according to the above we have that:

$$
\begin{aligned}
f(t) &= \int\!\!\int_{S^+(t)} \mu(\alpha,\beta)\hat{\gamma}_{\alpha\beta}u(t)d\alpha d\beta + \int\!\!\int_{S^-(t)} \mu(\alpha,\beta)\hat{\gamma}_{\alpha\beta}u(t)d\alpha d\beta \\
&= \int\!\!\int_{S^+(t)} \mu(\alpha,\beta)d\alpha d\beta - \int\!\!\int_{S^-(t)} \mu(\alpha,\beta)d\alpha d\beta
\end{aligned}
\tag{2.53}
$$

The following properties hold for the Preisach model.

(a) **Wiping-out Property** Each local input maximum wipes out the vertices of $L(t)$ whose α-coordinates are below this maximum and each local minimum wipes out the vertices whose β-coordinates are above this minimum. Namely, only the alternating series of dominant input extrema are stored by the Preisach model. All other input extrema are wiped out

(b) **Congruency Property** All minors hysteresis loops corresponding to back and forth variations of input between the same consecutive extremum values are congruent

39

What it follows next, is the determination of $\mu(\alpha, \beta)$. To determine $\mu(\alpha, \beta)$, the set of first-order transition (reversal) curves are needed. These curves can be experimentally found as follows. First the input $u(t)$ should be decreased to the value which is less than β_0 (a situation which is called negative saturation). Next the input is monotonically increased until it reaches some value α'. As the input increased, an ascending branch of a major loop is followed. This branch is also called as the limiting ascending branch, because usually there is no branch below it. The notation $f_{\alpha'}$ will be used for the output value on this branch which corresponds to the input value $u = \alpha'$. The first-order transition (reversal) curves are attached to the limiting ascending branch. Each of these curves is formed as the above monotonic increase of the input is followed by a sub-sequent monotonic decrease. The term first-order emphasize the fact that each of these curves is formed after the first reversal of input (see fig. 2.11).
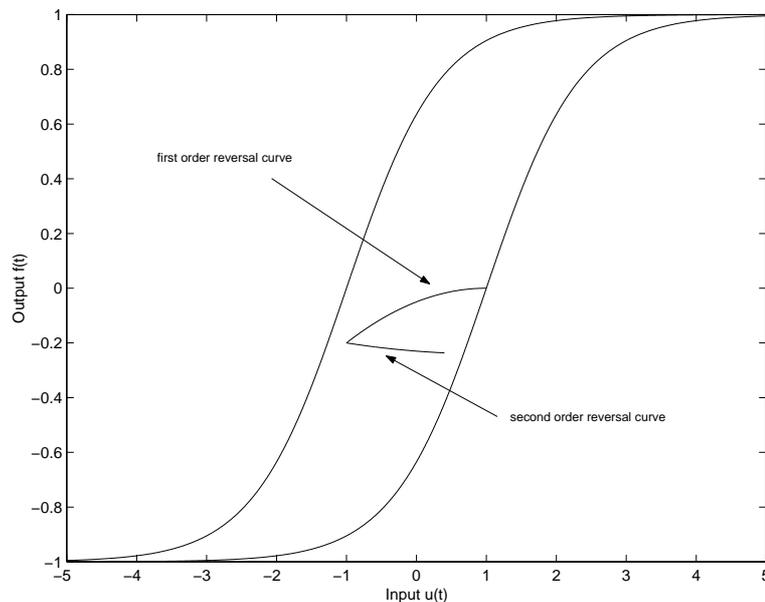

Figure 2.11: The First- and Second-order transition reversal curves

40

Defining the function:

$$F(\alpha', \beta') \triangleq \frac{1}{2}(f_{\alpha'} - f_{\alpha'\beta'}) \qquad (2.54)$$

after some calculations that contain differentiations, we have:

$$\mu(\alpha', \beta') = -\frac{\vartheta^2 F(\alpha', \beta')}{\vartheta\alpha'\vartheta\beta'} \qquad (2.55)$$

From (2.54), the above expression can be written in another equivalent form, which is:

$$\mu(\alpha', \beta') = \frac{1}{2}\frac{\vartheta^2 f_{\alpha'\beta'}}{\vartheta\alpha'\vartheta\beta'} \qquad (2.56)$$

These first-order curves can also be named as first-order decreasing transition curves. By almost repeating literally the previous reasoning a similar expression can be found by using the first-order increasing transition curves. It is important to note here that when the first-order transition curves are congruent the mirror symmetry of functions $F(\alpha, \beta)$ and $\mu(\alpha, \beta)$ with respect to the line $\alpha = -\beta$ is valid, i.e. :

$$F(-\beta, -\alpha) = F(\alpha, \beta)$$
$$\mu(-\beta, -\alpha) = \mu(\alpha, \beta)$$

respectively.

We next proceed to the formulation of the fundamental theorem that gives the necessary and sufficient conditions for the representation of actual hysteresis non-linearities by the Preisach model.

Representation Theorem

The wiping-out property and the congruency property constitute the necessary and sufficient conditions for a hysteresis nonlinearity to be represented by the Preisach model on the set of piecewise monotonic inputs.

We also mention that if the wiping-out and congruency properties are valid, then

it really does not matter which transition curves are used for the determination of $\mu(\alpha, \beta)$. If the properties are not valid the values of $\mu(\alpha, \beta)$ and the accuracy of the Preisach model will depend on a particular choice of transition curves employed for the determination of $\mu(\alpha, \beta)$.

The numerical implementation of the Preisach model which circumvents the evaluation of double integrals of (2.53) and the formula (2.56), where due to the differentiations may strongly amplify errors (noise) and are inherently presented in any experimental data, is:

$$
\begin{aligned}
f(t) = \quad - \quad & F(\alpha_0, \beta_0) + 2 \sum_{k=1}^{n(t)-1} \left[F(M_k, m_{k-1}) - F(M_k, m_k) \right] \\
+ \quad & 2 \left[F(M_n, m_{n-1}) - F(M_n, u(t)) \right]
\end{aligned}
\tag{2.57}
$$

where the above expression has been derived for monotonically decreasing input that is, the final link of interface $L(t)$ is a vertical one. For the case where we have monotonically increasing input :

$$
\begin{aligned}
f(t) = \quad - \quad & F(\alpha_0, \beta_0) + 2 \sum_{k=1}^{n(t)-1} \left[F(M_k, m_{k-1}) - F(M_k, m_k) \right] \\
+ \quad & 2 \left[F(u(t), m_{n-1}) \right]
\end{aligned}
\tag{2.58}
$$

The function $F(\alpha, \beta)$ is related to experimentally measured first-order transition curves by the formula (2.58), M_k and m_k correspond to the maximum and minimum values of the input respectively on the k-vertex of the Preisach phase plane and $n(t)$ represents the number of the dominant extremum values and it is a function of time due to the wiping-out property of the Preisach model and this number may change with the time.

The above discussed model is called *classical Preisach model* and constitutes the base to understand in depth the following modified models and that because, the classical model has some intrinsic limitations which are:

1. The C.P. (classical Preisach) model describes hysteresis nonlinearities which exhibit congruency of minor loops formed for the same reversal values of input. In fact the actual hysteresis nonlinearities deviate from this property.

2. The C.P. model is static in nature and does not account for dynamic properties of hysteresis nonlinearities. For fast input variations these properties may be essential.

3. The C.P. model describes hysteresis nonlinearities with the wiping-out property, which means to the immediate formation of hysteresis loop after one cycle of back-and-forth variation of input between any two reversal values. However, experiments show that hysteresis loop formation is often preceded by some stabilization process which may require large number of cycles to achieve a stable minor loop.

4. The C.P. model deals only with scalar hysteresis nonlinearities. In many applications however, vector hysteresis[1] is encountered.

The model that we will describe now is referred to as the *Moving Preisach model.* We subdivide the triangle T (see fig. 2.12) into three sets $S_{u(t)}^{+}, R_{u(t)}, S_{u(t)}^{-}$ which are defined as:

$$(\alpha, \beta) \in S_{u(t)}^{+} \quad \text{if } \beta_0 \leq \alpha \leq u(t)$$
$$(\alpha, \beta) \in R_{u(t)} \quad \text{if } \beta_0 \leq \beta \leq u(t), \quad u(t) \leq \alpha \leq \alpha_0$$
$$(\alpha, \beta) \in S_{u(t)}^{-} \quad \text{if } u(t) \leq \beta \leq \alpha \leq \alpha_0$$

The Moving Preisach model is the following:

$$f(t) = \int \int_{R_{u(t)}} \mu(\alpha, \beta) \hat{\gamma}_{\alpha\beta} u(t) d\alpha d\beta + \frac{1}{2}(f_{u(t)}^{+} + f_{u(t)}^{-}) \qquad (2.59)$$

---

[1]Vector hysteresis is a vector nonlinearity with the property that past extremum values of input projections along all possible directions may affect future values of output.

43
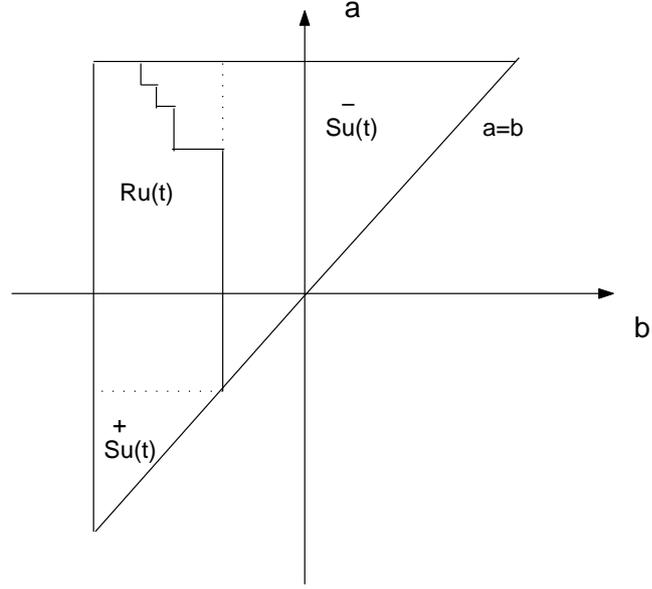
Figure 2.12: The Preisach phase plane for the Moving model

where $f_{u(t)}^+, f_{u(t)}^-$ is the output along the limiting ascending and the limiting descending branch respectively. In expression (2.59) the integration is performed not over the fixed limiting triangle T but over the rectangle $R_{u(t)}$which changes along with the input variations. The identification problem as before is in determining the μ-function by fitting the model (2.59) to some experimental data. To overcome this let introduce the function:

$$T(\alpha, \beta) \overset{\triangle}{=} f_{\beta}^- - f_{\alpha\beta} \tag{2.60}$$

where here assume that we started from the state of positive saturation and the input $u(t)$ is monotonically decreased. Consequently (2.60) is equal to the output increments between the limiting descending branch and first-order transition curves. After some calculations we conclude that:

$$\mu(\alpha, \beta) = -\frac{1}{2} \frac{\vartheta^2 T(\alpha, \beta)}{\vartheta \alpha \vartheta \beta} \tag{2.61}$$

44

and using (2.60) into (2.61) we have:

$$\mu(\alpha, \beta) = \frac{1}{2} \frac{\vartheta^2 f_{\alpha\beta}}{\vartheta\alpha\vartheta\beta} \qquad (2.62)$$

It must be cleared that the moving model (2.59) is equivalent to the classical Preisach model (2.52) as far as description of purely hysteretic behavior is concerned. More precisely, it is apparent that this equivalence holds only for input and output variations confined to the region enclosed by a major hysteresis loop. Outside this region, the C.P. model prescribes flat saturation values for output, while the moving model (2.59) prescribes the actual experimentally observed values $f^+_{u(t)}$ and $f^-_{u(t)}$ for the states of negative and positive saturation, respectively. For this reason, the wiping-out and congruency property of minor loop are valid for the moving model. Again, the numerical implementation is given by:

$$f(t) = 2 \sum_{k=1}^{n(t)} \left[T(M_{k+1}, m_k) - T(M_k, m_k) \right] + f^+_{u(t)} \qquad (2.63)$$

where (2.63) expresses explicitly the output $f(t)$ in terms of experimentally measured function T.

An another model which is a modified version of the classical Preisach model will be discussed now and is called as the *nonlinear or input dependent Preisach model*. The advantages of this model over the classical one are:

1. The congruency property of minor loops is relaxed and

2. The nonlinear model allows one to fit experimentally measured first and second-order reversal curves.

Since higher-order reversal curves are sandwiched between first- and second-order ones, it is reasonable to expect that the nonlinear model will be more accurate than the classical one. The nonlinear Preisach model can be mathematically defined as:

$$f(t) = \int\int_{R_{u(t)}} \mu(\alpha, \beta, u(t)) \hat{\gamma}_{\alpha\beta} u(t) d\alpha d\beta + \frac{f^+_{u(t)} + f^-_{u(t)}}{2} \qquad (2.64)$$

45

or using the geometric interpretation the equation (2.64) becomes:

$$f(t) = \int\int_{S(t)^+} \mu(\alpha, \beta, u(t)) d\alpha d\beta - \int\int_{S(t)^-} \mu(\alpha, \beta, u(t)) d\alpha d\beta \qquad (2.65)$$

It is clear that a new feature of this model in comparison with the moving one is the dependence of the function μ on the current value of input $u(t)$. This model has the following two properties:

(a) **Wiping-out Property** Only the alternating series of past dominant extrema M_k and m_k are stored by the nonlinear Preisach model.

(b) **Property of equal Vertical Chords** All minor loops resulting from back-and-forth variations between the same two consecutive extrema have equal vertical chords (output increments) for the same input values.

We should note that for two consecutive extrema of the input, let assume $u \in (u_+, u_-)$, the corresponding vertical chord does not depend on a particular past history preceding the formation of a minor loop.

Now, for the solution of the identification problem the sets of first and second-order reversal curves are required. Let assume that first the input is decreased to reach the negative saturation state and then monotonically increases until it reaches some value α. The first-order reversal curves are attached to the limiting ascending branch and they are formed when the above monotonic increase of $u(t)$ is followed by a subsequent decrease. The notation $f_{\alpha u}$ will be used for the output values on the first-order reversal curve. The second-order reversal curves are attached to the first-order reversal curves and they formed when the above monotonic decrease is followed by a monotonic increase (see fig. 2.11). Using $f_{\alpha\beta u}$ as the notation for the output values on the second-order reversal curve, we consider the function:

$$P(\alpha, \beta, u) \triangleq f_{\alpha u} - f_{\alpha\beta u} \qquad (2.66)$$

46

which (2.66) has the physical meaning of output increments between the first- and second-order reversal curves. It is clear that:

$$P(\alpha, u, u) = P(u, \beta, u) = P(u, u, u) = 0$$

After some calculations we obtain:

$$\mu(\alpha, \beta, u) = -\frac{1}{2} \frac{\vartheta^2 P(\alpha, \beta, u)}{\vartheta\alpha\vartheta\beta} \qquad (2.67)$$

and using (2.66) in (2.67) we finally obtain:

$$\mu(\alpha, \beta, u) = \frac{1}{2} \frac{\vartheta^2 f_{\alpha\beta u}}{\vartheta\alpha\vartheta\beta} \qquad (2.68)$$

Due to the mirror symmetry we also have that:

$$\mu(-\beta, -\alpha, -u) = \mu(\alpha, \beta, u)$$

The representation theorem for this model is the following one:

Representation Theorem

The wiping-out property and the property of equal vertical chords for minor loops constitute the necessary and sufficient conditions for the representation of a hysteresis nonlinearity by the nonlinear Preisach model on the set of piecewise monotonic inputs.

We must clear that the property of equal vertical chords is more general than the congruency property. Indeed, if comparable minor loops are congruent, then they have equal vertical chords. If now they have equal vertical chords, they are not necessarily congruent. Likewise, under the congruency condition the nonlinear Preisach model (2.64) or (2.65) coincides with the moving Preisach model (2.59). The numerical implementation of the nonlinear model is the following one:

$$f(t) = f_{u(t)}^- + \sum_{k=1}^{n(t)} \left[P(M_{k+1}, m_k, u(t)) \right] - P(M_k, m_k, u(t)) \qquad (2.69)$$

where $f^-_{u(t)}$ is the output value of a monotonic decrease input from some above α_0 (positive saturation) to α value of $u(t)$. The formula (2.69) computes output values by using input values, a set of second-order reversal curves and an input history which are all specified by the user. Other models, very similar to the above ones, can be found in [23].

Hysteresis modeling with the Preisach-model approach appears to be efficient as it better approaches the requirements of accuracy and adaptability; as a matter of fact, the possibility of including dynamic and mean field effects and the ability to be coupled with the numerical solutions of Maxwell equations justifies its large diffusion in many applications. From the other hand, the Preisach models incude integrodifferential operators, thus making them very complicated and it is still not clear how to fuse them into the controller design. However, we should mention that modeling a general type of hysteresis itself is still a research topic and the reader may refer to [24] for a recent view.

2.4   Summary

In this chapter mathematical models of non-smooth nonlinearities, such as friction, backlash and hysteresis, and their limitations have been presented. Identification and estimation techniques of the above models have also been presented when feasible. As it is cleared from this chapter, the better mathematical description we have the higher accuracy and better description of the nonlinearities achieved. From the other hand, the demand of high accuracy leads to complicated models that cannot be implemented sometimes into the controller design. The in depth knowledge of the above nonlinearities, the improvement of the nonlinear control theory that has been achieved in the last decade together with the advanced technology, orient the research to search for a unified method to face better the above nonlinearities.

As it will be cleared from the following chapter, which describes the basic principles of fault diagnosis, a succesful fault diagnosis is dependent from the model choice. The better model we have the more accurate fault diagnosis can be achieved.

Chapter 3

Fault Diagnosis

The detection and isolation of faults (diagnosis) is of a great importance in any engineering system. Such kind of systems can be a broad spectrum of human-made machinery, including industrial production facilities (oil refineries, steel mills, chemical plants, etc.) transportation vehicles (ships, airplanes, trains, etc.) and household devices (air conditioning equipment, refrigerators, washing machines, etc.). The early detection of the fault occurrence is critical in avoiding product deterioration, performance degradation, major damage to the machinery itself and damage to human health or even loss of lives. The quick and correct diagnosis of the faulty component then facilitates proper and optimal decisions on emergency and corrective actions and on repairs.

The traditional approaches to fault detection and diagnosis involve the limit checking of some variables or the application of redundant sensors (physical redundancy). More advanced methods rely on the spectral analysis of signals emanating from the machinery or on the comparison of the actual plant behavior to that expected on the basis of a mathematical model (analytical redundancy). The latter approach includes methods which are more deterministically framed (parity relations, observers) and those formulated more on a statistical basis (Kalman filtering and parameter

estimation). The boundaries between the various approaches are rather blurred and, lately, sever methods have been shown to be closely related to one another and even to produce identical results under broad conditions.

3.1 What is Fault Detection and Diagnosis

The detection and diagnosis of the faults in engineering systems are concerned whether they occur in the plant or in its measurement and control instruments. In the sequel what is meant by faults will be described and the tasks of detection and diagnosis will be specified.

In general, faults are deviations from the normal behavior in the plant or its instrumentation. The faults of interest belong to one of the following categories:

- Additive process faults.
  These are unknown inputs acting on the plant, which are normally zero and which, when present, cause a change in the plant outputs independent of the known inputs. Such faults best describe plant leaks, loads, etc.

- Multiplicative process faults.
  These are changes (abrupt or gradual) in some plant parameters. They cause changes in the plant ouputs which depend also on the magnitude of the known inputs. Such faults best describe the deterioration of plant equipment, such as surface contamination, clogging, or the partial or total loss of power.

- Sensor faults
  These are discrepancies between the measured and actual values of individual plant variables. These fautls are usually considered additive (independent of the measured magnitude), though some sensor faults (sticking or complete failure) may be better characterized as multiplicative.

- Actuator faults

  These are discrepancies between the input command of an actuator and its actual output. Actuator faults are usually handled as additive though, again, some kinds of them may be better characterized as multiplicative.

The sytems now that perform fault detection and diagnosis implement the following tasks:

- Faul detection, which means, the indication that something is going wrong in the monitored system.

- Fault isolation, which means, the determination of the exact location of the fault or in other words which component is faulty.

- Fault identification, which means, the determination of the magnitude of the fault.

The last two tasks together, that is, isolation and identification are referred to as fault diagnosis. While detection is absolutely necessary in any practical system and isolation is almost equally important, fault identification may not justify the extra effort it requires. For this reason, most practical systems contain only the fault detection and isolation tasks and are referred to as FDI systems. Most of the time, the fault detection and diagnosis activity takes place on-line, in real time. The two tasks can be performed either in parallel way or sequentially. In some systems, the detection tasks is running permanently while the diagnostic task is triggered only upon the detection of the presence of a fault.

Particularly, according to [26] in model-based fault detection and diagnosis (will be described later) the following conventions are usually adopted:

(i) It is assumed that the faults are not present initially in the system but arrive at some later time, The faults are generally described by a deterministic time-functions which are unknown.

(ii) Another deterministic and unknown inputs to the system are the additive disturbances. The distinction between additive faults and disturbances is subjective; the faults are those unknown inputs we wish to detect and isolate while disturbances are nuisances we wish to ignore.

(iii) The noise which emanates from the plant or from the sensors and actuators, is considered random with zero mean. Any nonzero mean is handled as a fault or a disturbance.

(iv) Modeling errors are discrepancies between the model (model parameters) and the true system. They are present ever since the origins of the system or due to the changes of the operating-point. They may be considered as multiplicative disturbances, in contrast to multiplicative faults which are also discrepancies between the model and the true system, but which we wish to detect.

The *detection performance* of the diagnostic technique is characterized by a number of important and quantifiable benchmarks which are:

- Fault sensitivity, that is, the ability of the technique to detect faults of reasonably small size.

- Reaction speed, that is, the ability of the technique to detect faults with reasonably small delay after their arrival.

- Robustness, that is, the ability of the technique to operate in the presence of noise, disturbances and modeling errors, with few false alarm[1]

It is remarkable to note here, that in the most cases, there are design trade-offs between the various properties described above. The *isolation performance*, that is, the ability of the diagnostic system to distinguish faults depend on the physical properties of the plant, on the size of the faults, noise, disturbances and modeling

---

[1]Erroneous fault detection

error, and on the design of the algorithm. Multiple simultaneous faults are, in general, more difficult to isolate them than single faults. Moreover, the interplay between faults and disturbances, noise, modeling error may lead to uncertain or incorrect isolation decisions. In addition, some faults may be non-isolable from one another because they act on the physical plant in an undistinguishable way.

3.2 Methods in Fault Detection and Diagnosis

The methods in fault detection and diagnosis (FDD), may be classified into two main categories: those that do not utilize a mathematical model of the plant and those that do.

3.2.1 Model-Free methods

These are the FDD methods that do not utilize a mathematical model of the plant and are:

Physical redundancy. In this method, multiple sensors are installed to measure the same physical quantity. Any serious discrepancy between the measurements indicates a sensor fault. With only two parallel sensors, fault isolation is not possible. With three sensors, a voting scheme can be formed in order to isolate the fault sensor. Physical redundancy involves extra hardware cost and extra weight, where in the latter consists a serious factor in aerospace applications.

Special sensors. These sensors are installed explicitly for detection and diagnosis purposes. They may be limit sensors (measuring e.g. temperature, pressure), which perform limit checking. Other special sensors may measure some faulty-indicating physical quantity, such as sound, vibration, etc.

Limit checking This approach is widely used in practice. Plant measurements are compared by computer to preset limits. Exceeding the threshold indicates a fault situation. While simple and straightforward, the limit checking approach has two

serious drawbacks:

- Since the plant variables may vary widely due to normal input variations, the thresholds need to be quite conservatively.

- The effect of a single component fault may propagate to many plant variables, setting off a confusing multitude of alarms and making isolation extremely difficult.

Spectrum analysis. Most plant variables exhibit a typical frequency spectrum under normal operating conditions; any deviation from this is an indication of abnormality. Certain types of faults may have a specific signature in the spectrum making thus isolation simpler.

Logic reasoning. Are techniques which are complementary to the discused above methods, in that they are aimed at evaluating the symptoms obtained by the detection (hardware and/or software). They are consist of trees of logical rules of the "IF symptom AND symptom THEN conclusion" type. Each conclusion in turn, can serve as a symptom in the next rule until to lead to a final conclusion.

3.2.2   Model-Based methods

Model-based FDD methods utilize an explicit mathematical model of the monitored plant. Their natural mathematical description is in the form of differential equations or equivalent transformed representations for the continuous-time model, while for the discrete-time in the form of difference equations or their transformed equivalents. Also, though most physical systems are nonlinear, their mathematical descriptions usually relies on linear approximations.

Most of the model-based FDD methods rely on the concept of *analytical redundancy*. In contrast with the physical redundancy, when measurements from parallel sensors are compared to each other, now sensory measurements are compared to analytically computed values of the respective variable. The resulting differences, called

residuals, are indicative of the presence of faults in the system. The generation of residuals needs to be followed by *residual evaluation*, in order to arrive at detection and isolation decisions. Schematically is depicted in fig. 3.1. Because of the presence
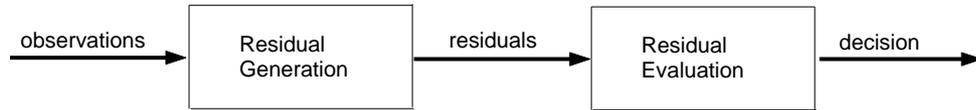

Figure 3.1: Stages of model-based fault detection and diagnosis.

of noise and modeling errors, the residuals are never zero, even if there is no fault. Hence, the detection decision requires testing the reiduals against thresholds, obtained empirically or by theoretical considerations. To facilitate fault isolation, the residual generators are usually designed for isolation enhanced residuals, exhibiting structural or directional properties. The isolation decisions then can be obtained in a structural (boolean) or directional (geometrical) framework, with or without the inclusion of statistical elements.

Robustness issues

The residuals which are generated to indicate faults may also react to the presence of noise, disturbances and modeling errors. Desensitizing the residuals to these sources is the most important aspect in the design of the detection and diagnosis algorithm. More precisely:

- To deal with the effects of noise, the residuals may be filtered and statistical techniques may be applied to their evaluation. In the case of not sufficient information concerning the statistical properties of the noise and the noise-transfer dynamics of the plant may complicate and hamper the overall procedure.

- Disturbance decoupling may be built into the design of the residuals genera-

tor, but it competes with the isolation enhancement for the available design freedom.

- Robustness in the face of modeling errors is the most fundamental problem in model-based FDD scheme. Several methods are available which usually rely on some sort of optimization. Unfortunately, this problem does not lend itself to easy solution and the known techniques are effective only under limited circumstances.

Residual Generation Techniques

The generation of residual signals is a central issue in model-based fault diagnosis. A rich variety of methods are available for residual generation and here will be discussed briefly some of the most common approaches. It must be pointed out that most residual generation approaches are applicable for both continuous and discrete models, however some approaches can only work for discrete models. For example, the parity relation approach is developed specially for discrete models although there have been some studies into the use of the parity relation approach for continuous models.

Considering the general cases, a system with all possible faults, i.e. sensor, actuator and process faults, according the methodology and the denomination of [27] can be described by the following state space model as:

$$
\begin{aligned}
\dot{x}(t) &= Ax(t) + Bu(t) + R_1 f(t) \\
y(t) &= Cx(t) + Du(t) + R_2 f(t)
\end{aligned}
\tag{3.1}
$$

where $f(t) \in \Re^g$ is a fault vector, each element $f_i(t)$ ($i = 1, 2, \cdots, g$) corresponds to a specific fault and is considered as unknown time function. The matrices R_1 and R_2 are known as fault entry matrices which represent the effects of faults on the system. The vector $u(t) \in \Re^r$ is the input to the actuator or measured actuation,

and the vector $y(t) \in \Re^m$ is the measured output, and both vectors are known for FDI purposes. The vector $x(t) \in \Re^n$ is the state vector and A, B, C, D are known system matrices with appropriate dimensions. What is following are some approaches that have been developed for the residual generation.

- *Observer-based approaches*

  The basic idea behind the observer or filter-based approaches is to estimate the outputs of the system from the measurements (or a subset of measurements) by using either Luenberger observer(s) in the deterministic setting or Kalman filter(s) in a stochastic setting. Then, the output estimation error (or innovations in the stochastic case), is used as a residual. It should be pointed out that we are interesting to estimate the outputs using an observer, while it is not necessary the estimation of the state vector. Therefore, a functional observer is suitable for this task. In practice, the order of the functional observer is less than the order of a state observer. It is desired to estimate a linear function of the state, i.e. $Lx(t)$, using a functional (or generalized) Luenberger observer with the following structure:

  $$
  \begin{aligned}
  \dot{z}(t) &= Fz(t) + Ky(t) + Ju(t) \\
  w(t) &= Gz(t) + Ry(t) + Su(t)
  \end{aligned} \tag{3.2}
  $$

  where $z(t) \in \Re^q$ is the state vector of this functional observer with F, K, J, R, G, S matrices with appropriate dimensions. The output $w(t)$ of this observer is said to be an estimate of $Lx(t)$, for the system described in (3.1), in an asymptotic sense if in the absence of faults:

  $$
  \lim_{t \to \infty} [w(t) - Lx(t)] = 0 \tag{3.3}
  $$

  To introduce a transformation matrix T, the observer in (3.2) will generate the estimate $Lx(t)$ in the asymptotic sense *if and only if* the following conditions

58

hold:

$$\left\{ \begin{array}{l} F \text{ has stable eigenvalues} \\[4pt] TA - FT = KC \\[4pt] J = TB - KD \\[4pt] RC + GT = L \\[4pt] S + RD = 0 \end{array} \right. \qquad (3.4)$$

The necessary and sufficient condition for the existence of the observer given by (3.2) for the system (3.1) is that the pair $(C,\ A)$ is observable. In order to generate residuals, we need to estimate the system output. If we assign:

$$L = C \qquad (3.5)$$

we have the output estimation as:

$$\hat{y}(t) = w(t) + Du(t) \qquad (3.6)$$

The residual vector $r(t)$ is defined as:

$$r(t) = Q[y(t) - \hat{y}(t)] = L_1 z(t) + L_2 y(t) + L_3 u(t) \qquad (3.7)$$

where:

$$
\begin{aligned}
L_1 &= -QG \\
L_2 &= Q - QR \\
L_3 &= -Q(S + D)
\end{aligned}
$$

Now, the residual generator based on generalized Luenberger is given by:

$$
\begin{aligned}
\dot{z}(t) &= Fz(t) + Ky(t) + Ju(t) \\
r(t) &= L_1 z(t) + L_2 y(t) + L_3 u(t)
\end{aligned}
\qquad (3.8)
$$

and the matrices in this equation should satisfy the following conditions:

$$
\begin{cases}
F & \text{has stable eigenvalues} \\
& TA - FT = KC \\
& J = TB - KD \\
& L_1 T + L_2 C = 0 \\
& L_3 + L_2 D = 0
\end{cases}
\tag{3.9}
$$

When we apply the residual generator described in (3.8) to the system described by (3.1), the residual will be:

$$
\begin{aligned}
\dot{e}(t) &= Fe(t) - TR_1 f(t) + KR_2 f(t) \\
r(t) &= L_1 e(t) + L_2 R_2 f(t)
\end{aligned}
\tag{3.10}
$$

where $e(t) = z(t) - Tx(t)$. It is obvious that the residual depends solely and totally on faults.

The simplest method is of the full order observer and in this case $(q = n)$ we have:

$$
\begin{aligned}
T &= I & L_1 &= QC \\
F &= A - KC & L_2 &= -Q \\
J &= B - KD & L_3 &= QD
\end{aligned}
$$

For any dynamic system, the observer-based residual generator always exists. This is because any input-output transfer function matrix has the observable realization. In other words, the output estimator always exists although a suitable state observer cannot always be designed. The minimal order q_0 of a functional observer satisfies the inequality:

$$
q_0 \leq \mu - 1
\tag{3.11}
$$

where μ is the observability index of the system which is defined as the mini-

mum number for which:

$$rank[C^T, (CA)^T, \cdots, (CA^\mu)^T] = n$$

For observable systems the observability index lies within the limits:

$$\frac{n}{m} \leq \mu \leq n - m + 1$$

Inequality (3.11) gives only the minimum possible order of a functional observer. Providing additional freedom in order to achieve the required diagnostic performance, the observer order is normally larger thna the minimum possible order.

To isolate the faults, the observer-based approaches can be used to design structured residual sets or fixed residual vectors. For sensor faults, such kind of design is straightforward. If it is required that a residual is sensitive to faults in all but one of the sensors, the observer used to generate this residual should be driven by outputs excluding that single sensor measurement. However, the design of a structured residual set for actuator fault isolation is more difficult. This problem can be solved via unknown input observers and eigenstructure assignement [27]. However, the isolation of actuators faults is not always possible.

- *Parity vector (relation) methods*

  The basic idea of the parity relation approach is to provide a proper check of the parity (consistency) of the measurements of the monitored system. The parity relations are rearranged direct input-output model equations, subjected to a linear dynamic transformation. The transformed residuals serve for detection and isolation. To begin with this problem, let consider the measurement of an n-dimensional vector using m sensors, as in [27]. The measurement equation is:

$$y(k) = Cx(k) + f(k) + \xi(k)$$

where $y(k) \in \Re^m$ the measurement vector, $x(k) \in \Re^n$ the state vector, $f(k)$ the vector of sensor faults , $\xi(k)$ the noise vector and C an $m \times n$ measurement matrix. Furthermore, the dimension of $y(k)$ is larger than the dimension of $x(k)$, that is:

$$m > n \quad \text{and} \quad rank(C) = n$$

Inconsistency in the measurement data is then a metric that can be used initially for detecting faults and, subsequently for fault isolation. For FDI purposes, the vector $y(k)$ can be combined into a set of linearly independent parity equations to generate the parity vector (residual):

$$r(k) = Vy(k)$$

To satisfy the usual requirement for a residual, that is zero-valued in the fault-free case, the matrix V should satisfy the condition:

$$VC = 0$$

Under this condition, the parity vector contains only information on the faults and noise:

$$r(k) = v_1[f_1(k) + \xi_1(k)] + \cdots + v_m[f_m(k) + \xi_m(k)] \qquad (3.12)$$

where v_i the ith column of V, $f_i(k)$ is the ith element of $f(k)$ which denotes the fault in the ith sensor. From (3.12) one can see that the parity vector is independent of the unmeasured state $x(k)$ and that contains information about the faults and the noise (uncertainty). Moreover, the parity space is spanned by the columns of V, i.e. the columns of V form a basis for this space. In addition, a fault in the ith sensor, implies a growth of the residual $r(k)$ in the direction v_i. The space $span\{V\}$ is called a "parity space". Then a fault detection decision function is defined as:

$$DFD(k) \triangleq r(k)^T r(k)$$

If a fault occurs in the sensors, $DFD(k)$ will be greater than a predetermined threshold. For the fault isolation decision another function is defined which is:

$$DFI_i(k) \triangleq v_i^T r(k) \; ; \; i \in \{1, \; 2, \cdots, m\}$$

For a given $r(k)$, a malfunctioning sensor is identified by computing the m values of $DFI_i(k)$. If $DFI_j(k)$ is the largest one of these values then the sensor that corresponds to $DFI_j(k)$ is the one which is most likely to have become faulty. In the parity space point of view, the columns of V define m distinct fault signature directions. After a fault has been declared, it can be isolated by comparing the orientation of the parity vector to each these signature directions. So, $DFI_i(k)$ is a measure of the correlation of the residual vector with fault signature directions. For a reliable isolation, the generalized angles between fault signature directions should be as large as possible, i.e., to make $v_i^T v_j \;\; (i \neq j)$ as small as possible. Thus, optimal fault isolation performance will be achieved when v_i determined by:

$$\begin{cases} min\{v_i^T v_j\} & i \neq j \;\; i, \; j \in \{1, \; 2, \cdots, \; m\} \\ max\{v_i^T v_i\} & i \in \{1, \; 2, \cdots, \; m\} \end{cases}$$

For the case $rank(C) = m < n$, redundancy relations are needed to be construct and can be done by collecting sensor outputs over a time interval. say $\{y(k-s), \; y(k-s+1), \cdots, y(k)\}$. This is known as "temporal" or "serial" redundancy. As in [27], we consider a system with the following discrete state space equations:

$$x(k+1) = Ax(k) + Bu(k) + R_1 f(k)$$
$$y(k) = Cx(k) + Du(k) + R_2 f(k) \tag{3.13}$$

where $y \in \Re^m$ the output vector, $x \in \Re^n$ the state vector, $u \in \Re^r$ the input vector, $f \in \Re^g$ the fault vector and $A, \; B, \; C, \; D, \; R_1, \; R_2$ real matrices with

compatible dimensions. Combining (3.13) from time instant $k - s$ to k yields the following redundant relations:

$$\underbrace{\begin{bmatrix} y(k-s) \\ y(k-s+1) \\ \vdots \\ y(k) \end{bmatrix}}_{Y(k)} - H \underbrace{\begin{bmatrix} u(k-s) \\ u(k-s+1) \\ \vdots \\ u(k) \end{bmatrix}}_{U(k)} = Wx(k-s) + M \underbrace{\begin{bmatrix} f(k-s) \\ f(k-s+1) \\ \vdots \\ f(k) \end{bmatrix}}_{F(k)}$$

(3.14)

or in a condensed form:

$$Y(k) - HU(k) = Wx(k-s) + MF(k) \tag{3.15}$$

with:

$$H = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{s-1}B & CA^{s-2}B & \cdots & D \end{bmatrix} \in \Re^{(s+1)m \times (s+1)r}$$

$$W = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^s \end{bmatrix} \in \Re^{(s+1)m \times n}$$

and the matrix M is constructed by replacing $\{D, \ B\}$ with $\{R_2, \ R_1\}$ in the matrix H. Then a residual signal can be defined as:

$$r(k) \overset{\triangle}{=} V[Y(k) - HU(k)] \tag{3.16}$$

where $V \in \Re^{p \times (s+1)m}$ and p the residual vector dimension. Eq.(3.16) is the computational form of a residual generator which shows the residual signal as a function of measured inputs and outputs of the monitored system. Using

64

(3.15) in (3.16) we obtain:

$$r(k) = VWx(k-s) + VMF(k) \qquad (3.17)$$

This is the evaluation format of the residual. Again to make the parity vector insensitive to system's inputs and states the following equation should hold:

$$VW = 0 \qquad (3.18)$$

and to satisfy the fault detectability condition, the matrix V should also satisfy the condition:

$$VM \neq 0 \qquad (3.19)$$

Once we have matrix V, the residual signal can be generated using (3.16). The residual generator design depends on solutions of (3.18). For an appropriately large s, it follows from the Cayley-Hamilton theorem that the solution of (3.18) exists and so the parity relation-based residual generator for fault detection does. It must be pointed out that the parity relation can also be constructed using a z-transformed input-output model.

The parity relation approach can be used to design structured residual set for fault isolation. As in observer-based approach, for isolating sensor faults is very straightforward. If we use c_i^T (the ith row of C) and y_i (the ith component of y) instead of C and y, the parity relation will contain only the ith sensor's output together with all the inputs. The residual generated by this relation is only sensitive to the fault in the ith sensor. For the actuator isolation problem, the structured residual set is more difficult to design and the isolation of actuator faults is not always possible.

- *FDI via Parameter Estimation*

  Parameter estimation is a natural approach to the detection and isolation of parametric (multiplicative) faults. This approach is based on the assumption

that the faults are reflected in the physical system parameters such friction, mass, resistance, etc. The basic idea of the detection method is that the parameters of the actual process are repeatedly estimated on-line using well-known estimation methods and the results are compared with the parameters the reference model obtained initially under the fault-free condition. Any substantial discrepancy indicated as a fault. This approach normally uses the I/O mathematical model of a system in the following form:

$$y(t) = f(P, u(t))$$

where P is the model coefficient vector which is directly related to physical parameters of the system. The function f can take both linear or non-linear foramts. The basic procedure that is followed for FDI purposes using parameter estimation approach has the following steps:

(1) Establish the process model using physical relations.

(2) Determine the relationship between model coefficients and process physical parameters.

(3) Estimate the normal model coefficients.

(4) Calculate the normal process physical parameters.

(5) Determine the parameter changes which occur for the various fault cases.

By carrying out the last step for known faults, a database of faults and their symptoms can be built up.

To generate residuals using this approach, an on-line parameter identification algorithm should be used. It is not easy however to achieve fault isolation using the parameter estimation method. This is because the parameters being identified are model parameters which cannot always be converted back to the system physical parameters. Moreover, this method is also more demanding

66

in terms of on-line computation and input excitaion requirements than the other methods described previously.

- *Kalman filter*

  The innovation (prediction error) of the Kalman filter can be used as a fault detection residual; its mean is zero if there is no fault (and disturbance) and becomes non-zero in the presence of faults. However, fault isolation is somewhat awkward with Kalman filter; one needs to run a bank of "matched filters", one for each suspected fault and for each possible arrival time, and check which filter output can be matched with the actual observation.

With their steady establisment in the past, reasearch attention has been devoted to the interconnection among these approaches, in particular, between parity relation and the other three approaches. Equivalence between them has been demonstrated form different viewpoints [30], [31]. Recently, [32], derived a one-to-one relationships among the design parameters and reveal that the real difference between these approaches lies in the fact the the on-line implementation form of parity relation approach is nonrecursive, while in the observer-based ones are implemented recursively. Making use of these results the design of residual generators can be carried out independent of the implementation form possibly used.We can use, for instance, the parity space approach for the residual generator design, then transform the parameters achieved to the parameters needed for the construction of a diagnostic observer and finally realize the diagnostic observer.

In the literature, all the above approaches described for the residual generation purposes, are referred to as *analytical approaches* that make use of quantitative models.

Other approaches that make use of qualitative models as well as approaches using computational intelligence techniques are referred to as *knowledge-based* approaches and will be described in the sequel.

- *Qualitative fault diagnosis*

  Since in most cases available, a priori knowledge about a process is hardly complete, or, even if this is the case, might be too complex to directly deal with, an approximation has to be made, so that models become inaccurate. Or measurements are subjected to noise. Consequently, deviations between the reality and its representation, i.e. modeling errors, are unavoidable. This method has been extensively applied in science and engineering, e.g. when nonlinear differential equation is linearized or a complex system is represented by a trained artificial neural network. These quantitative models are able to predict the system behavior precisely but more often inaccurately. Efforts have to be made through bringing more information (e.g. training data) to raise the accuracy of the prediction in the modeling stage, or through modeling error decoupling to reduce the influence of such errors when applying the models to fault diagnosis.

  Alternatively, incomplete knowledge can be treated via abstraction. Instead of the precise description by a quantitative model, a qualitative description of a process can be applied. By allowing the existence of a tolerance, the resolution of the representations is reduced, to emphasize primary distinctions and ignore unimportant or unknown details. Although this description is imprecise it is able to represent the system accurately. The qualitative approach, in contrast with quantitative one, requires only declarative information, e.g. the sign of variables, the tendencies of variables (increasing, decreasing or constant), order and/or relative magnitude, and hence can be robust with respect to uncertainty in a well defined sense. The qualitative approach is motivated by the following circumstances (see also [27]):

    - Faults cannot be reasonably described by analytical methods, e.g. a valve is blocked or a pipe is broken

- The on-line information available is not given by quantitative assesments of the current operating conditions, e.g. the water level is high cannot be unambiguously transformed into quantitative measurement data.

- If the system structure or parameters are not precisely known and diagnosis has to be based primarily on heuristic information (e.g. connection of symptoms and faults, process history, fault statistics etc.), no quantitative model can be set up.

According to the available information about a plant, there are several different possibilities to qualitatively represent the information of the dynamic process, each of which is associated with an appropriate simulation method. Basically, a qualitative simulation method should be responsible for retaining the accuracy of the represented system behavior, thus the fault detection approaches based on them could avoid false alarm. The representations that are relevant to the FDI approaches are:

- qualitative differential equations ([34])

- envelope behaviors ([35], [36])

- stochastic qualitative behaviors ([37], [33])

Main disadvantages of the qualitative approach emerge when there is a possibility of ambiguity, for example when are manipulated two or more declarative variables (the sum of a positive variable and a negative one can either be positive or negative), or because the qualitative models are relatively crude, usually cannot be used to detect soft faults as the diagnosis is symptom-based. Quantitative and qualitative approaches have a lot of complementary features and can be suitably combined together in order to increase the robustness of

the quantitative methods. This combination can also minimize the disadvantages of the two approaches. Hence, one of the aims in the future research on model-based FDI is to find the way to combine these two methods togeher to provide highly reliable diagnostic information.

- *FDI using Computational Intelligence Techniques*

In the case of fault diagnosis in complex systems, one is faced with the problem that no, or insufficiently accurate, mathematical models are available. The use of *knowledge-model-based* or *data-model-based* techniques, either in the framework of diagnosis expert systems or in combination with a human expert, is then the only feasible way to proceed.

Fuzzy logic in fault diagnosis

The second stage of model-based FDI, decision making, is a logic decision process that transforms quantitative knowledge (residual signals) into qualitative statements (faulty, normal, etc.). To outline the basic idea, let consider the case that the residual due to faults is also contaminated by noise and the effect of uncertainty due to incomplete de-coupling, so that the residual will be non-zero even in the absence of faults, i.e. the residual will fluctuate depending on the unknown time functions of the disturbances, noise and inputs of the process. Based upon this limitation, the problem is to make the correct decisions on the basis of uncertain information.

Contrary to the classical logic which allows a definite calssification of fixed values, the fuzzy logic offers a form for the description of tolerances. Fuzzy processing can be divided into essentially the following stages. In the first, the residuals are compared with membership functions which are often assumed to be of triangular shape. In the second stage, the lower of the two antecedent outputs is selected. Then the output of all rules is combined. Finally, the cen-

ter of gravity (or another averaging methods) is used to defuzzify the output and lead to the possibility of definite decision-making. The introduction of fuzzy logic can improve the decision-making, and in turn will provide reliable and sufficient FDI which are applicable for real industrial systems. One of the latest development in this area is the *fuzzy oberver based approach*. In [38], the fuzzy observer concept actually represents a set of analytical linear observers on whose ouputs a fuzzy fusion is performed based on Takaki-Sugeno fuzzy models. Using this approach a nonlinear dynamic system is described by a number of locally linearised models. For the fuzzy observer scheme the linear models are implemented in a bank of linear observers. The final state estimation is given by a fuzzy fusion of all local observer outputs. The difference between the measured output and the estimated output provides the residual for further diagnostic evaluation.

Neural Networks in fault diagnosis

In the past two decades, the techniques of artificial neural networks (ANN) are growing mature, as a data-driven method, which provides a totally new perspective to fault diagnosis. The ANN is hopeful approach to FDI, owing to its robustness and strong learning ability. The ability to learn means that, if a causal relationship exists between the output and input, the network will learn it. If sufficient internal nodes and internal layers are available, the network will also map any set of inputs to the corresponding output. The ultimate result is a network which will faithfully reproduce the desired output for the entire training set, including any noise. A very frequent application of ANNs for FDI purposes, is their use as classifiers with training data for each fault. However, the majority of the ANN-based FDI systems suffer form the lack of universality, the dilemma of stability and the long training time, due to

the localization of the algorithm itself. One of the latest development in this area is the *neural observer based approach*. In [39], [40] neural networks are used as nonlinear multi-input single-output models of ARMA type to set up different kinds of observer schemes. Thereby the neural networks replace the analytical models which are usually necessary for observer-based FDI. Two types of observer schemes are proposed by [39] for actuator, component and instrument fault detection: *the neural single observer scheme* and *the neural dedicated observer scheme*. Whilst the first one is driven by all process inputs and outputs the second one is only driven by the process inputs and the output of the component to be supervised. Therefore, the first scheme consists only of a single observer which is composed of a bank of multi-input single-output neural nets each estimating one output in contrast to the second scheme, which consists of a number of observers associated to each component of the plant. These neural observers in turn consist of a number of multi-input single-output neural nets each estimating one process output. In both cases the training is based on fault free process data reflecting the normal behavior. The residual evaluation part can then be performed by a well-known static multi-layer perceptron neural network.

It has to be mentioned also that a combination of the above intelligence techniques with the help of genetic algorithms can also be used in order to cope with the problem of nonlinear processes, lacking analytical knowledge and robustness issues. For example, fuzzy neural networks (FNNs) (see [41]) combine the advantage of fuzzy reasoning, which is the capability of handling uncertain and imprecise information, with the advantage of neural networks, which is the capability of learning from examples. Genetic algorithms from the other hand can be used in order to find for example, optimal neural structures.

Thanks to the rapid progress of nonlinear observer theory during the last decade, significant results in designing nonlinear residual generators have been achieved in recent five years. Nevertheless, a general theory for the solution of nonlinear FDI problems is still missing. Thus the development of nonlinear FDI approaches is one of the current FDI topics that are receiving much attention. Despite the difficulties, works on nonlinear systems have recently appeared [51]-[54].

3.3 Summary

This chapter has presented the basic priciples of FDI and especially of the model-based one. The FDI problem has been formalized in a uniform framework by presenting mathematical descriptions and definitions. The residual generator, which is identified as a central issue in model-based FDI, has been summarized in a generalized structure which can cover all residual generation methods. Other FDI methods such as computational intelligence techniques and qualitative modeling have been discussed briefly. In the following chapter, a novel approach for fault detection in mechanical systems is presented where friction nonlinearities are present. The precedening chapters including this one, have been constituted the base for the development of this novel approach.

Chapter 4

Fault detection in mechanical systems with friction phenomena: an on-line approximation approach.

In this chapter we present a novel approach to detect faults in mechanical systems with friction that perform linear motion. The basic module in the proposed architecture is an on-line approximator which is based on liner-in-the-weights neural network structures. To model the effects of friction, the dynamic LuGre model [9] is used. However, we don't assume knowledge of system nonlinearities. Furthermore, the friction internal state is not assumed to be available for measurement. The on-line approximator requires system's position and velocity as well as its input force. The performance of the developed fault detector is analyzed with respect to its robustness and sensitivity. Rigorous fault detectability conditions are also derived basing on the important results presented in [51].

74

4.1 Problem Formulation

Consider the linear motion of a mass m driven by an input force u:

$$m\ddot{x} + Kx + F = u \qquad (4.1)$$

where F represents the friction force, $K > 0$ denotes the spring constant, x the mass position, and $\dot{x}$ its velocity. To model the effects of friction, the dynamic LuGre model [9] is used:

$$F = \sigma_0 z + \sigma_1 \dot{z} + \dot{x} + \omega(x, \dot{x}, z) \qquad (4.2)$$

$$\dot{z} = -\alpha(\dot{x})|\dot{x}|z + \dot{x} \qquad (4.3)$$

where the friction internal state z describes the averaging deflection of the contact surfaces during the sticking phases and $\omega(x, \dot{x}, z)$ denotes a friction modelling error. We assume that $|\omega(x, \dot{x}, z)| \leq \bar{\omega}$, where $\bar{\omega} \geq 0$ is an *unknown but suitably small* (see Section 4.2) constant. Furthermore, the parameters σ_0, σ_1, σ_2 that appear in (4.2) are positive and are considered unknown, too. In [9], the function $\alpha(\dot{x})$ is given by

$$\alpha(\dot{x}) = \frac{\sigma_0}{f_c + (f_s - f_c)e^{-(\dot{x}/v_s)^2}}$$

where f_c is the Coulomb friction, f_s is the stiction force and v_s is the Stribeck velocity. It is apparent that $0 < \sigma_0/f_s \leq \alpha(\dot{x}) \leq \sigma_0/f_c$. In practice, $\alpha(\dot{x})$ depends on several factors such as material properties, temperature etc.

Defining $x_1 \triangleq x$, $x_2 \triangleq \dot{x}$, it follows that

$$\dot{x}_1 = x_2 \qquad (4.4)$$

$$\dot{x}_2 = -\alpha_1 x_2 + [\alpha_3 \alpha(x_2)|x_2| - \alpha_2]z + \alpha_4 u - \alpha_4 \omega(x_1, x_2, z) - \alpha_5 x_1 \qquad (4.5)$$

$$\dot{z} = -\alpha(x_2)|x_2|z + x_2 \qquad (4.6)$$

where $|x_2|\alpha(x_2) \geq 0$, $\alpha_1 \triangleq (\sigma_1 + \sigma_2)/m$, $\alpha_2 \triangleq \sigma_0/m$, $\alpha_3 \triangleq \sigma_1/m$, $\alpha_4 \triangleq 1/m$, and $\alpha_5 \triangleq K/m$.

For system (4.4)-(4.6), the following assumptions are introduced:

Assumption 1 *The state variables x_1 and x_2 are available for measurement.*

Assumption 2 *Let $\mathcal{U}$ be the class of piecewise continuous and bounded signals. Then, for any $u \in \mathcal{U}$ and any initial condition, the state trajectories x_1, x_2 are uniformly bounded.*

It is worth noting that the internal friction state z is **not** assumed to be available for measurement, and the function $\alpha(x_2)$ as well as the positive parameters α_i, $i = 1, \ldots, 5$ are considered unknown.

As $\alpha(x_2)|x_2| > 0$ and $\alpha(x_2)|x_2| = 0$ only when $x_2 = 0$, it follows immediately that the internal friction state z is input–to–state stable when x_2 is considered as input. Hence, there exists a pair of functions $\beta \in \mathcal{KL}$ and $\gamma \in \mathcal{K}$ (these functions are not assumed to be known) such that, for every essentially bounded input x_2, we have

$$|z(t, z_0, x_2)| \leq \beta(|z_0|, t) + \gamma(|x_2|), \quad \forall t \geq 0 \tag{4.7}$$

where $z(t, z_0, x_2)$ denotes the trajectory of (4.6) starting from z_0 at time $t = 0$ with input x_2.

The faults considered in this paper are modelled as additive perturbations (occurring at some unknown time instant T) $\Delta F_1(x_1, x_2, t)$ and $\Delta F_2(x_1, x_2, t)$ to the nominal F and K in (4.1), respectively. Then, after occurrence of a fault (i.e., for $t \geq T$), the dynamics of the systems becomes

$$\dot{x}_1 = x_2 \tag{4.8}$$
$$\dot{x}_2 = -\alpha_1 x_2 - \alpha_5 x_1 + \alpha_4 u + [\alpha_3 \alpha(x_2)|x_2| - \alpha_2] z - \alpha_4 \omega(x_1, x_2, z)$$
$$+ [(\alpha_3 \alpha(x_2)|x_2| - \alpha_2) z - \alpha_1 x_2 - \alpha_4 \omega(x_1, x_2, z)] \Delta F_1 - \alpha_5 x_1 \Delta F_2 \tag{4.9}$$
$$\dot{z} = -\alpha(x_2)|x_2| z + x_2 \tag{4.10}$$

The following further assumption is introduced (no multiple faults are considered in this paper).

Assumption 3 *Only one single fault may occur at a given time T.*

To end this section on the problem formulation, we would like to emphasize that the additive perturbations $\Delta F_1(x_1, x_2, t)$ and $\Delta F_2(x_1, x_2, t)$ to the nominal values of F and K, respectively, reflect variations in the normal forces in contact, temperature changes and material wear, as well as spring's stiffening and relaxation phenomena. These malfunctions are typically encountered for instance in actuators installed in harsh plant environments (see also the DAMADICS benchmark problem presented in Chapter 5, where the problem of detecting faults in a sugar plant actuator is addressed).

4.2 Nominal System On-line Approximation

In this section, we present an on–line approximation scheme for the nominal system presented in Section 4.1. The approximator's output will serve as the residual signal for fault detection. In this respect, as will be seen later on, a key role is played by the functional approximation scheme that in this work is implemented by one–hidden–layer neural structures with a linear output layer. In the following, the basic properties of such a class of neural approximators will be briefly reported for the sake of completeness.

More specifically, the considered class of neural approximators can be characterized as

$$y^\top = W^\top S(v) \tag{4.11}$$

where $v \in \Re^{n_2}$ and $y \in \Re^{n_1}$ denote the approximator input and output, respectively, W is an L-dimensional vector of synaptic weights, and $S(v)$ is a $L \times n_1$ matrix of regressor terms.

The regressor terms may contain high order connections of sigmoidal functions [47], radial basis functions (RBFs) with fixed centers and widths [44], [48], [46], shifted sigmoids [42], [43], thus forming High Order Neural Networks (HONNs), RBFs and Shifted Sigmoidal Neural Networks, respectively.

An important and well known property shared by the aforementioned neural approximating structures is the following (see also the references above):

Density Property. *For every continuous function* $f(v) : \Re^{n_2} \to \Re^{n_1}$, *there exist an integer* L *and optimal weight values* $W^\star$ *such that for every* $\epsilon > 0$

$$\sup_{v \in \Omega} |f(v)^\top - W^{\star^\top} S(v)| \leq \epsilon$$

where $\Omega \subset \Re^{n_2}$ *is a given compact set.*

In other terms, if the number of regressor terms L is sufficiently large, then there exist a weight vector $W^\star$ such that $W^{\star^\top} S(v)$ can approximate $f(v)^\top$ to any degree of accuracy, in a given compact set. This property allows us to focus on linear–in–the–weights neural networks (LNN for short) without loss of generality in terms of approximation error. This, in turn, will make it easier to prove basic system properties like stability and robustness. However, it is also important to mention that, under suitable assumptions, neural networks are characterized by other interesting properties related to their appproximating capabilities (see the basic work [49] and also the recent paper [50], where an extensive discussion on such properties in the nonlinear optimal control context is reported).

Now, let us consider the following estimator

$$\dot{\hat{x}}_1 \;=\; x_2 + k_1 \tilde{\xi}_1 - \hat{\alpha}_5 \tilde{\xi}_2 \tag{4.12}$$

$$\dot{\hat{x}}_2 \;=\; -\hat{\alpha}_1 \hat{x}_2 - \hat{\alpha}_5 \hat{x}_1 + \hat{\alpha}_4 u + k_2 \tilde{\xi}_2 - \phi \,, \tag{4.13}$$

where $k_1, k_2 > 0$ denote design constants, ϕ denotes a function that will be defined later on, $\tilde{\xi}_1$, $\tilde{\xi}_2$ represent the state estimation errors defined as

$$\tilde{\xi}_i \;\triangleq\; x_i - \hat{x}_i, \quad i = 1, 2 \,, \tag{4.14}$$

and $\hat{\alpha}_i$, for $i = 1, 4, 5$ are parameters to be updated on line. Then, from (4.4),(4.5), and (4.12)-(4.14), it follows that

$$\dot{\tilde{\xi}}_1 \;=\; -k_1 \tilde{\xi}_1 + \hat{\alpha}_5 \tilde{\xi}_2 \tag{4.15}$$

$$\dot{\tilde{\xi}}_2 \;=\; -\alpha_5 \tilde{\xi}_1 - (k_2 + \alpha_1)\tilde{\xi}_2 + \tilde{\alpha}_1 \hat{x}_2 + \tilde{\alpha}_5 \hat{x}_1 - \tilde{\alpha}_4 u + [\alpha_3 \alpha(x_2)|x_2| - \alpha_2]z -$$
$$-\alpha_4 \omega(x_1, x_2, z) + \phi \tag{4.16}$$

where $\tilde{\alpha}_i \;\triangleq\; \hat{\alpha}_i - \alpha_i$, for $i = 1, 4, 5$.

The objective is to design an adaptive structure for ϕ such as to guarantee the boundedness of the estimation errors and of all internal variables in front of the unknown friction terms entering the dynamics of the system via the unmeasurable state variable z and of the modelling uncertainty.

In this connection, we introduce the following form for function ϕ in (4.13):

$$\phi \;\triangleq\; -[|x_2|\hat{w}_1^\top S_1(x_2, |x_2|) + \hat{w}_2^\top S_2(|x_2|) + |x_2|\hat{\epsilon}_1 + \hat{b}_1]sigm(\tilde{\xi}_2) \tag{4.17}$$

where $sigm(\cdot)$ denotes a sigmoidal smooth approximation of the signum function $sgn(x)$, the terms $\hat{w}_1^\top S_1(x_2, |x_2|)$ and $\hat{w}_2^\top S_2(|x_2|)$ denote neural approximators of the form (4.11) and $\hat{\epsilon}_1, \hat{b}_1$ are further parameters to be updated on line. The reasons

motivating structure (4.17) for term ϕ will be clear after the proof of Theorem 1 stated later on.

It is worth noting that

$$sigm(x) = sgn(x) + \varepsilon_s(x) \tag{4.18}$$

where the error $\varepsilon_s(x)$ satisfies $|\varepsilon_s(x)| \leq 1$ (as is well know, $sigm(\cdot)$ can be shaped to make $\varepsilon_s(x)$ as small as desired). Hence, using (4.17) and (4.18), we obtain that (4.17) can be rewritten as

$$\phi = -[|x_2|\hat{w}_1^\top S_1(x_2, |x_2|) + \hat{w}_2^\top S_2(|x_2|) + |x_2|\hat{\epsilon}_1 + \hat{b}_1](sgn(\tilde{\xi}_2) + \varepsilon_s(\tilde{\xi}_2)) \tag{4.19}$$

The parameters appearing in (4.12),(4.13), and (4.19) are provided by the following adaptive laws:

$$
\begin{aligned}
\dot{\hat{w}}_1 &= P_a\{|\tilde{\xi}_2||x_2|S_1(x_2, |x_2|)\} \\
\dot{\hat{w}}_2 &= P_b\{|\tilde{\xi}_2|S_2(|x_2|)\} \\
\dot{\hat{\epsilon}}_1 &= P_c\{|\tilde{\xi}_2||x_2|\} \\
\dot{\hat{b}}_1 &= P_d\{|\tilde{\xi}_2|\} \\
\dot{\hat{\alpha}}_1 &= P_e\{-\hat{x}_2\tilde{\xi}_2\} \\
\dot{\hat{\alpha}}_4 &= P_f\{u\tilde{\xi}_2\} \\
\dot{\hat{\alpha}}_5 &= P_g\{-\tilde{\xi}_2(\tilde{\xi}_1 + \hat{x}_1)\}
\end{aligned} \tag{4.20}
$$

where P_a, P_b, P_c, P_d, P_e, P_f and P_g denote the projection operators with respect to the convex sets $\mathcal{W}_a \triangleq \{\hat{w}_1 \in \Re^{L_1} : |\hat{w}_1| \leq M_a\}$, $\mathcal{W}_b \triangleq \{\hat{w}_2 \in \Re^{L_2} : |\hat{w}_2| \leq M_b\}$, $\mathcal{W}_c \triangleq \{\hat{\epsilon}_1 \in \Re : |\hat{\epsilon}_1| \leq M_c\}$, $\mathcal{W}_d \triangleq \{\hat{b}_1 \in \Re : |\hat{b}_1| \leq M_d\}$, $\mathcal{W}_e \triangleq \{\hat{\alpha}_1 \in \Re : 0 < \hat{\alpha}_1 \leq M_e\}$, $\mathcal{W}_f \triangleq \{\hat{\alpha}_4 \in \Re : 0 < \hat{\alpha}_4 \leq M_f\}$, $\mathcal{W}_g \triangleq \{\hat{\alpha}_5 \in \Re : 0 < \hat{\alpha}_5 \leq M_g\}$, where M_a, M_b, M_c, M_d, M_e, M_f, M_g are suitably large positive scalars (the definition of the projection operation with respect to a convex set can be found, for instance, in [45]).

Now, we are able to state and prove the following basic theorem:

Theorem 1 *Consider the system (4.4)-(4.6). There exists a choice of the scalars M_a, M_b, M_c, M_d, M_e, M_f, M_g defining the projection sets $\mathcal{W}_a$, $\mathcal{W}_b$, $\mathcal{W}_c$, $\mathcal{W}_d$, $\mathcal{W}_e$, $\mathcal{W}_f$, $\mathcal{W}_g$ and of the initial conditions of the estimated parameters in the adaptive laws (4.20) such that the on-line approximator (4.12),(4.13), (4.19) together with the update laws (4.20) guarantee the uniform ultimate boundedness of $\tilde{\xi}_i$, $i = 1, 2$ with respect to the sets*

$$
\begin{aligned}
\Xi_1 &= \left\{ \tilde{\xi}_1 \in \Re \middle| |\tilde{\xi}_1| \leq \frac{\bar{\bar{\Phi}} M_g}{(k_2 + \alpha_1)k_1} \right\} \\
\Xi_2 &= \left\{ \tilde{\xi}_2 \in \Re \middle| |\tilde{\xi}_2| \leq \frac{|\Phi|}{k_2 + \alpha_1} \leq \frac{\bar{\bar{\Phi}}}{k_2 + \alpha_1} \right\}
\end{aligned}
$$

where $\bar{\bar{\Phi}} > |\Phi| = [|x_2|\hat{w}_1^\top S_1(x_2, |x_2|) + \hat{w}_2^\top S_2(|x_2|) + |x_2|\hat{\epsilon}_1 + \hat{b}_1]$ is a suitable positive scalar, as well as the boundedness of all parameter estimates $\hat{w}_i$, $i = 1, 2$, $\hat{\epsilon}_1$, $\hat{b}_1$, $\hat{\alpha}_1$, $\hat{\alpha}_4$ and $\hat{\alpha}_5$.

Proof: Consider the Lyapunov function candidate:

$$
V \triangleq \frac{1}{2}\tilde{\xi}_1^2 + \frac{1}{2}\tilde{\xi}_2^2 + \frac{1}{2}\tilde{w}_1^\top \tilde{w}_1 + \frac{1}{2}\tilde{w}_2^\top \tilde{w}_2 + \frac{1}{2}\tilde{\epsilon}_1^2 + \frac{1}{2}\tilde{b}_1^2 + \frac{1}{2}\tilde{\alpha}_1^2 + \frac{1}{2}\tilde{\alpha}_4^2 + \frac{1}{2}\tilde{\alpha}_5^2 \quad (4.21)
$$

Differentiating V with respect to time we obtain

$$
\begin{aligned}
\dot{V} &= -k_1\tilde{\xi}_1^2 + \hat{\alpha}_5\tilde{\xi}_1\tilde{\xi}_2 - (k_2 + \alpha_1)\tilde{\xi}_2^2 - \alpha_5\tilde{\xi}_1\tilde{\xi}_2 + \tilde{\xi}_2(\alpha_3\alpha(x_2)|x_2| - \alpha_2)z + \phi\tilde{\xi}_2 + \tilde{w}_1^\top \dot{\tilde{w}}_1 + \tilde{w}_2^\top \dot{\tilde{w}}_2 \\
&\quad -\tilde{\xi}_2\alpha_4\omega(x_1, x_2, z) + \tilde{\xi}_2\tilde{\alpha}_1\hat{x}_2 + \tilde{\xi}_2\tilde{\alpha}_5\hat{x}_1 - \tilde{\xi}_2\tilde{\alpha}_4 u + \tilde{\alpha}_1\dot{\hat{\alpha}}_1 + \tilde{\alpha}_4\dot{\hat{\alpha}}_4 + \tilde{\alpha}_5\dot{\hat{\alpha}}_5 + \tilde{\epsilon}_1\dot{\hat{\epsilon}}_1 + \tilde{b}_1\dot{\hat{b}}_1 \\
&\leq -k_1\tilde{\xi}_1^2 - (k_2 + \alpha_1)\tilde{\xi}_2^2 + \tilde{\xi}_1\tilde{\xi}_2\tilde{\alpha}_5 + |\tilde{\xi}_2||x_2|\alpha_3\alpha(x_2)|z| + \alpha_2|\tilde{\xi}_2||z| + \tilde{\xi}_2\phi + \tilde{w}_1^\top \dot{\tilde{w}}_1 + \tilde{w}_2^\top \dot{\tilde{w}}_2 \\
&\quad +|\tilde{\xi}_2||\alpha_4|\omega(x_1, x_2, z)| + \tilde{\xi}_2\tilde{\alpha}_1\hat{x}_2 + \tilde{\xi}_2\tilde{\alpha}_5\hat{x}_1 - \tilde{\xi}_2\tilde{\alpha}_4 u + \tilde{\alpha}_1\dot{\hat{\alpha}}_1 + \tilde{\alpha}_4\dot{\hat{\alpha}}_4 + \tilde{\alpha}_5\dot{\hat{\alpha}}_5 + \tilde{\epsilon}_1\dot{\hat{\epsilon}}_1 + \tilde{b}_1\dot{\hat{b}}_1
\end{aligned}
$$

$$(4.22)$$

From bounds (4.7) and $|\omega(x, \dot{x}, z)| \leq \bar{\omega}$, and introducing a positive scalar d_0 such that $0 \leq \beta(|z_0|, t) \leq d_0$, it follows that

$$
\dot{V} \leq -k_1\tilde{\xi}_1^2 - (k_2 + \alpha_1)\tilde{\xi}_2^2 + |\tilde{\xi}_2||x_2|\alpha_3\alpha(x_2)\Big[\beta(|z_0|, t) + \gamma(|x_2|)\Big] + \alpha_2|\tilde{\xi}_2|\Big[\beta(|z_0|, t) + \gamma(|x_2|)\Big]
$$

$$+\tilde{\xi}_2\phi + \tilde{w}_1^\top \dot{\tilde{w}}_1 + \tilde{w}_2^\top \dot{\tilde{w}}_2 + |\tilde{\xi}_2|\alpha_4\bar{\omega} + \tilde{\xi}_2\tilde{\alpha}_1\hat{x}_2 + \tilde{\xi}_2\tilde{\alpha}_5(\tilde{\xi}_1 + \hat{x}_1) - \tilde{\xi}_2\tilde{\alpha}_4 u + \tilde{\alpha}_1\dot{\tilde{\alpha}}_1 + \tilde{\alpha}_4\dot{\tilde{\alpha}}_4$$

$$+\tilde{\alpha}_5\dot{\tilde{\alpha}}_5 + \tilde{\epsilon}_1\dot{\tilde{\epsilon}}_1 + \tilde{b}_1\dot{\tilde{b}}_1$$

$$= -k_1\tilde{\xi}_1^2 - (k_2 + \alpha_1)\tilde{\xi}_2^2 + |\tilde{\xi}_2||x_2|\beta(|z_0|,t)\alpha_3\alpha(x_2) + |\tilde{\xi}_2||x_2|\alpha_3\alpha(x_2)\gamma(|x_2|) + \alpha_2|\tilde{\xi}_2|\beta(|z_0|,t)$$

$$+\alpha_2|\tilde{\xi}_2|\gamma(|x_2|) + \tilde{\xi}_2\phi + \tilde{w}_1^\top\dot{\tilde{w}}_1 + \tilde{w}_2^\top\dot{\tilde{w}}_2 + |\tilde{\xi}_2|\alpha_4\bar{\omega} + \tilde{\xi}_2\tilde{\alpha}_1\hat{x}_2 + \tilde{\xi}_2\tilde{\alpha}_5(\tilde{\xi}_1 + \hat{x}_1) - \tilde{\xi}_2\tilde{\alpha}_4 u$$

$$+\tilde{\alpha}_1\dot{\tilde{\alpha}}_1 + \tilde{\alpha}_4\dot{\tilde{\alpha}}_4 + \tilde{\alpha}_5\dot{\tilde{\alpha}}_5 + \tilde{\epsilon}_1\dot{\tilde{\epsilon}}_1 + \tilde{b}_1\dot{\tilde{b}}_1$$

$$\leq -k_1\tilde{\xi}_1^2 - (k_2 + \alpha_1)\tilde{\xi}_2^2 + |\tilde{\xi}_2||x_2|\Big[\alpha_3\alpha(x_2)\big(d_0 + \gamma(|x_2|)\big)\Big] + |\tilde{\xi}_2|(\alpha_2 d_0 + \alpha_4\bar{\omega}) + |\tilde{\xi}_2|\alpha_2\gamma(|x_2|)$$

$$+\tilde{\xi}_2\phi + \tilde{w}_1^\top\dot{\tilde{w}}_1 + \tilde{w}_2^\top\dot{\tilde{w}}_2 + \tilde{\xi}_2\tilde{\alpha}_1\hat{x}_2 + \tilde{\xi}_2\tilde{\alpha}_5(\tilde{\xi}_1 + \hat{x}_1) - \tilde{\xi}_2\tilde{\alpha}_4 u + \tilde{\alpha}_1\dot{\tilde{\alpha}}_1 + \tilde{\alpha}_4\dot{\tilde{\alpha}}_4 + \tilde{\alpha}_5\dot{\tilde{\alpha}}_5$$

$$+\tilde{\epsilon}_1\dot{\tilde{\epsilon}}_1 + \tilde{b}_1\dot{\tilde{b}}_1$$

The idea now is to approximate on line the unknown nonlinear terms $\alpha_3\alpha(x_2)(d_0 + \gamma(|x_2|))$ and $\alpha_2\gamma(|x_2|)$ by suitable neural approximators. More specifically, it turns out that there exist continuous functions $\varepsilon_1(x_2)$, $\varepsilon_2(x_2)$ (denoting the approximation errors) and constant but unknown weight vectors $w_1^\star$, $w_2^\star$, such that

$$\begin{aligned} \alpha_3\alpha(x_2)\left(d_0 + \gamma(|x_2|)\right) &= w_1^{\star\top} S_1(x_2, |x_2|) + \varepsilon_1(x_2) \\ \alpha_2\gamma(|x_2|) &= w_2^{\star\top} S_2(|x_2|) + \varepsilon_2(x_2) \end{aligned} \tag{4.23}$$

From the density property, it also follows that, on a generic compact set $\Omega \subset \Re$, the approximation errors can be suitably bounded as $|\varepsilon_1(x_2)| \leq \epsilon_1$ and $|\varepsilon_2(x_2)| \leq \epsilon_2$, where $\epsilon_1 > 0$, $\epsilon_2 > 0$.

Now, letting $k_3^\star \triangleq \alpha_2 d_0 + \alpha_4\bar{\omega}$ and using (4.23), we have

$$\begin{aligned} \dot{V} \leq{}& -k_1\tilde{\xi}_1^2 - (k_2 + \alpha_1)\tilde{\xi}_2^2 + |\tilde{\xi}_2||x_2|\Big[w_1^{\star\top} S_1(x_2, |x_2|) + \varepsilon_1(x_2)\Big] + k_3^\star|\tilde{\xi}_2| \\ &+ |\tilde{\xi}_2|\Big[w_2^{\star\top} S_2(|x_2|) + \varepsilon_2(x_2)\Big] + \tilde{\xi}_2\phi + \tilde{w}_1^\top\dot{\tilde{w}}_1 + \tilde{w}_2^\top\dot{\tilde{w}}_2 + \tilde{\xi}_2\tilde{\alpha}_1\hat{x}_2 + \tilde{\xi}_2\tilde{\alpha}_5(\tilde{\xi}_1 + \hat{x}_1) \\ &- \tilde{\xi}_2\tilde{\alpha}_4 u + \tilde{\alpha}_1\dot{\tilde{\alpha}}_1 + \tilde{\alpha}_4\dot{\tilde{\alpha}}_4 + \tilde{\alpha}_5\dot{\tilde{\alpha}}_5 + \tilde{\epsilon}_1\dot{\tilde{\epsilon}}_1 + \tilde{b}_1\dot{\tilde{b}}_1 \\ ={}& -k_1\tilde{\xi}_1^2 - (k_2 + \alpha_1)\tilde{\xi}_2^2 + |\tilde{\xi}_2||x_2|w_1^{\star\top} S_1(x_2, |x_2|) + k_3^\star|\tilde{\xi}_2| + |\tilde{\xi}_2|w_2^{\star\top} S_2(|x_2|) \\ &+ |\tilde{\xi}_2|\Big[|x_2|\varepsilon_1(x_2) + \varepsilon_2(x_2)\Big] + \tilde{\xi}_2\phi + \tilde{w}_1^\top\dot{\tilde{w}}_1 + \tilde{w}_2^\top\dot{\tilde{w}}_2 + \tilde{\xi}_2\tilde{\alpha}_1\hat{x}_2 + \tilde{\xi}_2\tilde{\alpha}_5(\tilde{\xi}_1 + \hat{x}_1) \end{aligned}$$

$$-\tilde{\xi}_2 \tilde{\alpha}_4 u + \tilde{\alpha}_1 \dot{\hat{\alpha}}_1 + \tilde{\alpha}_4 \dot{\hat{\alpha}}_4 + \tilde{\alpha}_5 \dot{\hat{\alpha}}_5 + \tilde{\epsilon}_1 \dot{\hat{\epsilon}}_1 + \tilde{b}_1 \dot{\hat{b}}_1$$

$$
\begin{aligned}
\leq \quad & -k_1 \tilde{\xi}_1^2 - (k_2 + \alpha_1) \tilde{\xi}_2^2 + |\tilde{\xi}_2| |x_2| w_1^{\star \top} S_1(x_2, |x_2|) + |\tilde{\xi}_2| w_2^{\star \top} S_2(|x_2|) + |\tilde{\xi}_2| \left[|x_2| \epsilon_1 + k_3^\star + \epsilon_2 \right] \\
& + \tilde{\xi}_2 \phi + \tilde{w}_1^\top \dot{\hat{w}}_1 + \tilde{w}_2^\top \dot{\hat{w}}_2 + \tilde{\xi}_2 \tilde{\alpha}_1 \hat{x}_2 + \tilde{\xi}_2 \tilde{\alpha}_5 (\tilde{\xi}_1 + \hat{x}_1) - \tilde{\xi}_2 \tilde{\alpha}_4 u + \tilde{\alpha}_1 \dot{\hat{\alpha}}_1 + \tilde{\alpha}_4 \dot{\hat{\alpha}}_4 + \tilde{\alpha}_5 \dot{\hat{\alpha}}_5 \\
& + \tilde{\epsilon}_1 \dot{\hat{\epsilon}}_1 + \tilde{b}_1 \dot{\hat{b}}_1
\end{aligned}
$$

After adding and subtracting the terms $|\tilde{\xi}_2| |x_2| \hat{w}_1^\top S_1(x_2, |x_2|)$ and $|\tilde{\xi}_2| \hat{w}_2^\top S_2(|x_2|)$, we obtain

$$
\begin{aligned}
\dot{V} \;=\; & -k_1 \tilde{\xi}_1^2 - (k_2 + \alpha_1) \tilde{\xi}_2^2 + |\tilde{\xi}_2| |x_2| w_1^{\star \top} S_1(x_2, |x_2|) + |\tilde{\xi}_2| |x_2| \hat{w}_1^\top S_1(x_2, |x_2|) \\
& - |\tilde{\xi}_2| |x_2| \hat{w}_1^\top S_1(x_2, |x_2|) + |\tilde{\xi}_2| w_2^{\star \top} S_2(|x_2|) + |\tilde{\xi}_2| \hat{w}_2^\top S_2(|x_2|) - |\tilde{\xi}_2| \hat{w}_2^\top S_2(|x_2|) \\
& + |\tilde{\xi}_2| \left[|x_2| \epsilon_1 + b_1 \right] + \tilde{\xi}_2 \phi + \tilde{w}_1^\top \dot{\hat{w}}_1 + \tilde{w}_2^\top \dot{\hat{w}}_2 + \tilde{\xi}_2 \tilde{\alpha}_1 \hat{x}_2 + \tilde{\xi}_2 \tilde{\alpha}_5 (\tilde{\xi}_1 + \hat{x}_1) \\
& - \tilde{\xi}_2 \tilde{\alpha}_4 u + \tilde{\alpha}_1 \dot{\hat{\alpha}}_1 + \tilde{\alpha}_4 \dot{\hat{\alpha}}_4 + \tilde{\alpha}_5 \dot{\hat{\alpha}}_5 + \tilde{\epsilon}_1 \dot{\hat{\epsilon}}_1 + \tilde{b}_1 \dot{\hat{b}}_1 \\
\;=\; & -k_1 \tilde{\xi}_1^2 - (k_2 + \alpha_1) \tilde{\xi}_2^2 - |\tilde{\xi}_2| |x_2| \tilde{w}_1^\top S_1(x_2, |x_2|) + |\tilde{\xi}_2| |x_2| \hat{w}_1^\top S_1(x_2, |x_2|) - |\tilde{\xi}_2| \tilde{w}_2^\top S_2(|x_2|) \\
& + |\tilde{\xi}_2| \hat{w}_2^\top S_2(|x_2|) + |\tilde{\xi}_2| \left[|x_2| \epsilon_1 + b_1 \right] + \tilde{\xi}_2 \phi + \tilde{w}_1^\top \dot{\hat{w}}_1 + \tilde{w}_2^\top \dot{\hat{w}}_2 \\
& + \tilde{\xi}_2 \tilde{\alpha}_1 \hat{x}_2 + \tilde{\xi}_2 \tilde{\alpha}_5 (\tilde{\xi}_1 + \hat{x}_1) - \tilde{\xi}_2 \tilde{\alpha}_4 u + \tilde{\alpha}_1 \dot{\hat{\alpha}}_1 + \tilde{\alpha}_4 \dot{\hat{\alpha}}_4 + \tilde{\alpha}_5 \dot{\hat{\alpha}}_5 + \tilde{\epsilon}_1 \dot{\hat{\epsilon}}_1 + \tilde{b}_1 \dot{\hat{b}}_1
\end{aligned}
$$

Using (4.19), (4.20) and defining

$$\Phi \overset{\triangle}{=} [|x_2| \hat{w}_1^\top S_1(x_2, |x_2|) + \hat{w}_2^\top S_2(|x_2|) + |x_2| \hat{\epsilon}_1 + \hat{b}_1]$$

we obtain

$$
\begin{aligned}
\dot{V} \;\leq\; & -k_1 \tilde{\xi}_1^2 - (k_2 + \alpha_1) \tilde{\xi}_2^2 - \Phi \varepsilon_s(\tilde{\xi}_2) \tilde{\xi}_2 \\
\leq\; & -(k_2 + \alpha_1) \tilde{\xi}_2^2 + |\Phi \varepsilon_s(\tilde{\xi}_2) \tilde{\xi}_2| \\
\leq\; & -(k_2 + \alpha_1) \tilde{\xi}_2^2 + |\Phi| |\tilde{\xi}_2| = -|\tilde{\xi}_2| \left[(k_2 + \alpha_1) |\tilde{\xi}_2| - |\Phi| \right] \\
\leq\; & 0
\end{aligned}
$$

provided that

$$|\tilde{\xi}_2| > \frac{|\Phi|}{k_2 + \alpha_1}$$

83

Hence $\tilde{\xi}_2$ is uniformly ultimately bounded with respect to the set

$$\Xi_2 = \left\{\tilde{\xi}_2 \in \Re \middle| |\tilde{\xi}_2| \leq \frac{|\Phi|}{k_2 + \alpha_1}\right\}$$

Notice that

$$|\Phi| \leq |x_2||\hat{w}_1||S_1(x_2, |x_2|)| + |\hat{w}_2||S_2(|x_2|)| + |x_2||\hat{\epsilon}_1| + |\hat{b}_1|$$

Moreover, $|S_1(x_2, |x_2|)|$, $|S_2(|x_2|)|$ are bounded too (let s_1, s_2 denote their known upper bounds). Then, according to (4.20) and denoting by $\bar{x}_2$ the known upper bound on the velocity (see Assumption 2), we have

$$|\Phi| \leq \bar{\Phi} \stackrel{\triangle}{=} \bar{x}_2 M_a s_1 + M_b s_2 + \bar{x}_2 M_c + M_d$$

Furthermore, since

$$\dot{\tilde{\xi}}_1 = -k_1 \tilde{\xi}_1 + \hat{\alpha}_5 \tilde{\xi}_2$$

we obtain

$$\tilde{\xi}_1(t) = |\tilde{\xi}_1(0)|e^{-k_1 t} + \int_0^t e^{-k_1(t-\tau)}\hat{\alpha}_5 \tilde{\xi}_2(\tau)d\tau$$

Thus,

$$|\tilde{\xi}_1(t)| \leq |\tilde{\xi}_1(0)|e^{-k_1 t} + M_g \int_0^t e^{-k_1(t-\tau)}|\tilde{\xi}_2(\tau)|d\tau$$

which finally becomes

$$|\tilde{\xi}_1(t)| \leq |\tilde{\xi}_1(0)|e^{-k_1 t} + \frac{\bar{\Phi}M_g}{(k_2 + \alpha_1)k_1}(1 - e^{-k_1 t})$$

Hence, $\tilde{\xi}_1(t)$ and $\tilde{\xi}_2(t)$ are uniformly ultimately bounded with respect to the sets

$$\Xi_1 = \left\{\tilde{\xi}_1 \in \Re \middle| |\tilde{\xi}_1| \leq \frac{\bar{\Phi}M_g}{(k_2 + \alpha_1)k_1}\right\}$$

$$\Xi_2 = \left\{\tilde{\xi}_2 \in \Re \middle| |\tilde{\xi}_2| \leq \frac{|\Phi|}{k_2 + \alpha_1} \leq \frac{\bar{\Phi}}{k_2 + \alpha_1}\right\}$$

Now, if the sets $\mathcal{W}_a$, $\mathcal{W}_b$, $\mathcal{W}_c$, $\mathcal{W}_d$, $\mathcal{W}_e$, $\mathcal{W}_f$, $\mathcal{W}_g$ are chosen in such a way that $w_1^\star$, $\hat{w}_1(0) \in \mathcal{W}_\alpha$, $w_2^\star$, $\hat{w}_2(0) \in \mathcal{W}_b$, ϵ_1, $\hat{\epsilon}_1(0) \in \mathcal{W}_c$, b_1, $\hat{b}_1(0) \in \mathcal{W}_d$, α_1, $\hat{\alpha}_1(0) \in \mathcal{W}_e$,

α_4, $\hat{\alpha}_4(0) \in \mathcal{W}_f$ and α_5, $\hat{\alpha}_5(0) \in \mathcal{W}_g$, where $\hat{w}_1(0)$, $\hat{w}_2(0)$, $\hat{\epsilon}_1(0)$, $\hat{b}_1(0)$, $\hat{\alpha}_1(0)$, $\hat{\alpha}_4(0)$ and $\hat{\alpha}_5(0)$ denote the initial values of $\hat{w}_1$, $\hat{w}_2$, $\hat{\epsilon}_1$, $\hat{b}_1$, $\hat{\alpha}_1$, $\hat{\alpha}_4$ and $\hat{\alpha}_5$, respectively, then the use of the projection modification to the update laws (4.20) guarantees the boundedness of $\hat{w}_1$, $\hat{w}_2$, $\hat{\epsilon}_1$, $\hat{b}_1$, $\hat{\alpha}_1$, $\hat{\alpha}_4$ and $\hat{\alpha}_5$, thus ending the proof of the theorem.

$\square$

Remark 1. It is worth noting that the magnitude of sets Ξ_1 and Ξ_2 depends on several factors and in general it is not easy to ascertain a clear way as to how reduce this magnitude that cannot be made arbitrarily small. In this respect, a key role is played by the modeling uncertainty and the number L of regressor terms on one side, and by the design constants k_1 and k_2 on the other. In case of significant modeling uncertainties, several issues come up. First of all, the projection algorithm may not be able to guarantee the boundedness of the signals that appear in (4.20). More specifically, the projection modification requires knowledge of the upper bounds on the norms of the unknown weights, b_1, ϵ_1, etc. If, for instance, we have $|b_1| > M_d$ due to the modeling error, the parameter $\hat{b}_1$ may drift to infinity since there is no guarantee that $\hat{b}_1$ will be bounded. Moreover, in the presence of large modeling error, large values of the design constants k_1, k_2 are needed in order to maintain the sets Ξ_1, Ξ_2 reasonably small. However, large values of k_1, k_2 may give rise to high gain feedback which in turn leads to instability. These important issues deserve further investigation.

As it has been shown by Theorem 1, the velocity error $\tilde{\xi}_2$ is guaranteed to be uniformly ultimately bounded with respect to the set Ξ_2. Consequently (the assumptions of the theorem being satisfied), if $\tilde{\xi}_2(0) \notin \Xi_2$, then there exists a finite time $T_0 > 0$ such that $\tilde{\xi}_2(t) \in \Xi_2$, $\forall\, t \geq T_0$. This important, though not surprising, result can be proved as follows. Recall that

$$\dot{V} \leq -(k_2 + \alpha_1)\tilde{\xi}_2^2 + \Phi\varepsilon_s(\tilde{\xi}_2)\tilde{\xi}_2 \leq -k_2\tilde{\xi}_2^2 + \Phi\varepsilon_s(\tilde{\xi}_2)\tilde{\xi}_2 \leq -k_2|\tilde{\xi}_2|^2 + |\Phi||\tilde{\xi}_2| \quad (4.24)$$

Moreover, recall also that $\dot{V} \leq 0$, whenever $|\tilde{\xi}_2| > \frac{|\Phi|}{k_2 + \alpha_1}$. Then integrating (4.24) from 0 to T_0, we obtain

$$V(T_0) - V(0) \leq \int_0^{T_0} (-k_2|\tilde{\xi}_2|^2 + |\Phi||\tilde{\xi}_2|)dt$$

which becomes (according also to the Lyapunov function definition):

$$\frac{1}{2}|\tilde{\xi}_2|^2 \leq V(T_0) \leq V(0) + \int_0^{T_0} (-k_2|\tilde{\xi}_2|^2 + |\Phi||\tilde{\xi}_2|)dt$$

It is sufficient to show that there exist a finite T_0 such that

$$V(0) + \int_0^{T_0} (-k_2|\tilde{\xi}_2|^2 + |\Phi||\tilde{\xi}_2|)dt \leq \frac{1}{2}\left(\frac{|\Phi|}{k_2 + \alpha_1}\right)^2$$

or

$$V(0) \leq \left[\int_0^{T_0} (k_2|\tilde{\xi}_2|^2 - |\Phi||\tilde{\xi}_2|)dt\right] + \frac{1}{2}\left(\frac{|\Phi|}{k_2 + \alpha_1}\right)^2 \tag{4.25}$$

Because we assumed before that $\tilde{\xi}_2(0) \notin \Xi_2$, using (4.24) we have that

$$-k_2|\tilde{\xi}_2|^2 + |\Phi||\tilde{\xi}_2| < 0 \quad \forall t \in [0, T_0]$$

Hence,

$$\int_0^{T_0} (-k_2|\tilde{\xi}_2|^2 + |\Phi||\tilde{\xi}_2|)dt < 0 \tag{4.26}$$

Observe that by (4.26) the term in brackets in (4.25) is positive. Moreover, define

$$\Gamma(T_0) = \int_0^{T_0} (k_2|\tilde{\xi}_2|^2 - |\Phi||\tilde{\xi}_2|)dt$$

Then

$$\frac{d\Gamma(T_0)}{dT_0} = k_2|\tilde{\xi}_2(T_0)|^2 - |\Phi(T_0)||\tilde{\xi}_2(T_0)| > 0$$

Since $\Gamma(T_0)$ is monotonically increasing and positive, there exists a finite time T_0 which satisfies (4.25). Hence,

$$|\tilde{\xi}_2(T_0)| \leq \frac{|\Phi(T_0)|}{k_2 + \alpha_1}$$

which contradicts what we assumed before. If on the other hand we restrict $\tilde{\xi}_2(0) \in \Xi_2$ then $T_0 = 0$.

Summing up, if Theorem 1 holds true, we have shown that the velocity error enters the set Ξ_2 in finite time. If such a set can be made sufficiently small, this result can be exploited in the framework of fault detection as will be seen in the next section.

4.3 Fault Detectability Analysis

In the previous section, the robustness properties of the on-line approximation scheme prior to the occurrence of a possible fault have been analyzed. Now, assume that conditions under which Theorem 1 holds true are satisfied and, accordingly, let T_0 to have the same meaning as before, that is, let it denote the time instant at which the nominal trajectory of the velocity error $\tilde{\xi}_2(t)$ enters the set Ξ_2 and never leaves it $\forall\, t \geq T_0$.

Now, consider the occurrence of a fault at time T in which case the dynamics of the system is described by Eqs. (4.8)-(4.10). The following further assumption is needed.

Assumption 4 *The time instant T of fault occurrence satisfies $T > T_0$.*

Clearly, if Assumption 4 is satisfied, no false alarm is generated prior to the occurrence of a fault, provided that $|\tilde{\xi}_2(t)|$ serves as the residual signal and the threshold is selected as

$$\rho \triangleq \frac{|\Phi|.}{k_2} > \frac{|\Phi|.}{k_2 + \alpha_1} \tag{4.27}$$

Remark 2. The threshold function that is used, is the conservative $\rho = \frac{|\Phi|}{k_2}$ instead of the uniform bound that appears in the definition of Ξ_2, because the parameter α_1 is considered unknown.

The decision on the occurrence of a fault is being made when the residual signal exceeds the threshold, i.e.

$$\text{a fault occurred if} \quad \exists\, T > T_0 \quad \text{such that} \quad |\tilde{\xi}_2(T)| > \rho$$

This decision criterion reflects the very intuitive fact that the fault to be detectable should be big enough to make the residual exceeding the threshold. In this respect, it is thus very important to address the issue of *fault detectability*.

The analysis in this section is deeply inspired by the basic work by Polycarpou and Trunov [51] with two differences: the state vector is not assumed to be completely available for measurement and the on-line approximator operates $\forall\, t \geq 0$ and not only after detection of a fault.

After occurrence of a fault (i.e., $t \geq T$), from Eqs. (4.9), (4.13), and (4.14) it follows that

$$
\begin{aligned}
\dot{\tilde{\xi}}_2 \;=\; & -\alpha_5 \tilde{\xi}_1 - (k_2 + \alpha_1)\tilde{\xi}_2 + (\alpha_3 \alpha(x_2)|x_2| - \alpha_2)z - \alpha_4 \omega(x_1, x_2, z) + \phi + \tilde{\alpha}_1 \hat{x}_2 \\
& + \tilde{\alpha}_5 \hat{x}_1 - \tilde{\alpha}_4 u + ((\alpha_3 \alpha(x_2)|x_2| - \alpha_2)z - \alpha_1 x_2 - \alpha_4 \omega(x_1, x_2, z)) \cdot \Delta F_1(x_1, x_2, t) \\
& - \alpha_5 x_1 \Delta F_2(x_1, x_2, t)
\end{aligned}
$$

$$(4.28)$$

(Recall that we assume a single fault scenario and thus ΔF_1 and ΔF_2 cannot be simultaneously different from zero.) Moreover, let

$$A \;\triangleq\; -\alpha_5 \tilde{\xi}_1 + [\alpha_3 \alpha(x_2)|x_2| - \alpha_2]z - \alpha_4 \omega(x_1, x_2, z) + \phi + \tilde{\alpha}_1 \hat{x}_2 + \tilde{\alpha}_5 \hat{x}_1 - \tilde{\alpha}_4 u$$

$$(4.29)$$

$$B_1 \;\triangleq\; [(\alpha_3 \alpha(x_2)|x_2| - \alpha_2)z - \alpha_1 x_2 - \alpha_4 \omega(x_1, x_2, z)] \cdot \Delta F_1(x_1, x_2, t) \qquad (4.30)$$

$$B_2 \;\triangleq\; -\alpha_5 x_1 \Delta F_2(x_1, x_2, t) \qquad (4.31)$$

According to [51], the detectability analysis can be performed in both the abrupt and the incipient fault cases. Specifically, an incipient time–profile for the fault can be characterized by a multiplicative term $(1 - e^{-\pi(t-T)})$, where $\pi > 0$ is an unknown constant that represents the rate evolution of the fault. In case $\pi = \infty$ the fault becomes an abrupt one.

The following simple result (analogous to the one presented in [51] for generic nonlinear systems) characterizes, in an implicit way, the set of faults that can be detected using the previously defined threshold.

Theorem 2 *Assume that fault* $\Delta F_i(x_1, x_2, t)$*, for* $i = 1$ *or* $i = 2$ *occurs at time* T*. If there exists a time interval* $[T + t_1, T + t_2]$*, with* $t_2 > t_1 \geq 0$*, such that*

$$\left| \int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)}(1 - e^{-\pi(\tau-T)}) B_i d\tau \right|$$

$$\geq \rho + \rho e^{-(k_2+\alpha_1)(t_2-t_1)} + \left| \int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)} A d\tau \right| \quad (4.32)$$

with $\rho = \frac{|\Phi|}{k_2}$ *and* Φ *as defined in Theorem 1. Then the fault is detected at time* $t = t_2$*.*

Proof. For any $t_2 > t_1$ the solution of (4.28) using (4.29) and (4.30) or (4.31), is given by:

$$\tilde{\xi}_2(T + t_2) = e^{-(k_2+\alpha_1)(T+t_2-T-t_1)} \tilde{\xi}_2(T + t_1) + \int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)} A d\tau$$

$$+ \int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)}(1 - e^{-\pi(\tau-T)}) B_i d\tau$$

Using the triangle inequality and $|\tilde{\xi}_2(T + t_1)| \leq \rho = \frac{|\Phi|}{k_2}$, we obtain:

$$|\tilde{\xi}_2(T + t_2)| \geq -\rho e^{-(k_2+\alpha_1)(t_2-t_1)} - \left| \int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)} A d\tau \right|$$

$$+ \left| \int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)}(1 - e^{-\pi(\tau-T)}) B_i d\tau \right|$$

where if the fault function is such that (4.32) is satisfied, then we obtain that $|\tilde{\xi}_2(T + t_2)| \geq \rho$, which implies that the fault will be detected.

$\square$

Estimation of the detection time

One of the most important characteristics in any fault diagnosis scheme is the time (detection time) required between the occurrence and the detection of a fault. Early detection (i.e, small detection time), is crucial to prohibit the possibly catastrophic consequences of a fault.

The following result (the proof is inspired again by [51]) gives an upper bound on the detection time for abrupt and incipient faults.

Theorem 3 *Assume that Theorem 2 holds. Moreover, suppose that there exist lower bounds $B_{mi} \leq B_i$, $i = 1, 2$ and an upper bound $\bar{A} > A$ such that, for $i = 1, 2$, we have*

$$B_{mi} > \bar{A} + |\Phi|, \ \forall t \in [T + t_1 \ T + t_d]$$

Then:

(a) **incipient faults**: *an upper bound t_d^+ on the detection time t_d is given by the solution of the algebraic equation*

$$g_i(t_d^+, k_2 + \alpha_1) - \left[g_i(t_1, k_2 + \alpha_1) + |\Phi| - \bar{A} \right] e^{-(k_2+\alpha_1)(t_d^+ - t_1)} = \bar{A} + |\Phi| \quad (4.33)$$

where

$$g_i(t, k_2 + \alpha_1) = \frac{B_{mi}}{k_2 + \alpha_1 - \pi_i}(k_2 + \alpha_1 - \pi_i - (k_2 + \alpha_1)e^{-\pi_i t})) \quad (4.34)$$

(b) **abrupt faults**: *an upper bound t_d^+ on the detection time t_d is given by*

$$t_d^+ = \frac{1}{k_2 + \alpha_1} ln \left[\frac{B_{mi} - \bar{A} + |\Phi|}{B_{mi} - \bar{A} - |\Phi|} \right] + t_1 \quad (4.35)$$

Furthermore in general, t_d^+ decreases monotonically as k_2 increases.

Proof. (a) As $\bar{A}$ is an upper bound on A, the following inequality holds:

$$\left| \int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)} A d\tau \right| \leq \bar{A} \int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)} d\tau = \frac{\bar{A}}{k_2+\alpha_1}(1-e^{-(k_2+\alpha_1)(t_2-t_1)})$$

(4.36)

Similarly, as B_{mi} is a lower bound on B_i, we have:

$$\left| \int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)}(1-e^{-\pi_i(\tau-T)}) B_i d\tau \right| \geq B_{mi} \int_{t_1}^{t_2} e^{-(k_2+\alpha_1)(t_2-\tau)}(1-e^{-\pi_i\tau}) d\tau$$

$$= \frac{B_{mi}}{k_2+\alpha_1} - \frac{B_{mi}}{k_2+\alpha_1}e^{-(k_2+\alpha_1)(t_2-t_1)} - \frac{B_{mi}}{k_2+\alpha_1-\pi_i}e^{-\pi_i t_2}$$

$$+ \frac{B_{mi}}{k_2+\alpha_1-\pi_i}e^{-(k_2+\alpha_1)t_2+(k_2+\alpha_1-\pi_i)t_1}$$

$$= \frac{B_{mi}}{(k_2+\alpha_1)(k_2+\alpha_1-\pi_i)}\left[(k_2+\alpha_1-\pi_i) - (k_2+\alpha_1-\pi_i)e^{-(k_2+\alpha_1)(t_2-t_1)} - (k_2+\alpha_1)e^{-\pi_i t_2}\right.$$

$$\left. + (k_2+\alpha_1)e^{-(k_2+\alpha_1)(t_2-t_1)}e^{-\pi_i t_1}\right]$$

$$= \frac{B_{mi}}{(k_2+\alpha_1)(k_2+\alpha_1-\pi_i)}\left[(k_2+\alpha_1-\pi_i) - (k_2+\alpha_1)e^{-\pi_i t_2}\right]$$

$$- \frac{B_{mi}}{(k_2+\alpha_1)(k_2+\alpha_1-\pi_i)}\left[(k_2+\alpha_1-\pi_i) - (k_2+\alpha_1)e^{-\pi_i t_1}\right]e^{-(k_2+\alpha_1)(t_2-t_1)}$$

$$= \frac{g_i(t_2, k_2+\alpha_1)}{k_2+\alpha_1} - \frac{g_i(t_1, k_2+\alpha_1)}{k_2+\alpha_1}e^{-(k_2+\alpha_1)(t_2-t_1)}$$

(4.37)

Hence, using (4.36) and (4.37), it follows that the detectability condition (4.32) becomes

$$\frac{g_i(t_2, k_2+\alpha_1)}{k_2+\alpha_1} - \frac{g_i(t_1, k_2+\alpha_1)}{k_2+\alpha_1}e^{-(k_2+\alpha_1)(t_2-t_1)}$$

$$\geq \rho e^{-(k_2+\alpha_1)(t_2-t_1)} + \rho + \frac{\bar{A}}{k_2+\alpha_1} - \frac{\bar{A}}{k_2+\alpha_1}e^{-(k_2+\alpha_1)(t_2-t_1)} \quad (4.38)$$

An upper bound on the detection time can thus be obtained by solving with respect to the unknown t_d^+ the algebraic equation

$$\frac{g_i(t_d^+, k_2+\alpha_1)}{k_2+\alpha_1} - \frac{g_i(t_1, k_2+\alpha_1)}{k_2+\alpha_1}e^{-(k_2+\alpha_1)(t_d^+-t_1)} = \rho e^{-(k_2+\alpha_1)(t_d^+-t_1)} + \rho$$

$$+ \frac{\bar{A}}{k_2+\alpha_1} - \frac{\bar{A}}{k_2+\alpha_1}e^{-(k_2+\alpha_1)(t_d^+-t_1)}$$

or, equivalently,

$$g_i(t_d^+, k_2 + \alpha_1) - \left[g_i(t_1, k_2 + \alpha_1) + |\Phi| - \bar{A}\right]e^{-(k_2+\alpha_1)(t_d^+ - t_1)} = \bar{A} + |\Phi|$$

thus proving (4.33).

(b) Letting $\pi_i \to \infty$, it follows that (4.33) becomes

$$B_{mi} - \left(B_{mi} + |\Phi| - \bar{A}\right)e^{-(k_2+\alpha_1)(t_d - t_1)} = \bar{A} + |\Phi|$$

and hence

$$e^{(k_2+\alpha_1)(t_d - t_1)} = \frac{B_{mi} - \bar{A} + |\Phi|}{B_{mi} - \bar{A} - |\Phi|}$$

thus obtaining

$$t_d = \frac{1}{k_2 + \alpha_1} ln \left[\frac{B_{mi} - \bar{A} + |\Phi|}{B_{mi} - \bar{A} - |\Phi|}\right] + t_1$$

which proves (4.35).

Finally, let us show that t_d^+ decreases monotonically as k_2 increases. From (4.38) and recalling that $\rho = \frac{|\Phi|}{k_2}$, we have

$$
\begin{aligned}
g_i(t_d, k_2 + \alpha_1) - g_i(t_1, k_2 + \alpha_1)e^{-(k_2+\alpha_1)(t_d - t_1)} &\geq \bar{A} + \frac{|\Phi|(k_2 + \alpha_1}{k_2} + (\bar{A} - |\Phi|)e^{-(k_2+\alpha_1)(t_d - t_1)} \\
&> \bar{A} + |\Phi| + (\bar{A} - |\Phi|)e^{-(k_2+\alpha_1)(t_d - t_1)} \quad (4.39)
\end{aligned}
$$

It is useful to introduce the quantities

$$
\begin{aligned}
f &\triangleq g_i(t_d, k_2 + \alpha_1) - g_i(t_1, k_2 + \alpha_1)e^{-(k_2+\alpha_1)(t_d - t_1)} \\
z &\triangleq \bar{A} + |\Phi| + (\bar{A} - |\Phi|)e^{-(k_2+\alpha_1)(t_d - t_1)}
\end{aligned}
$$

The partial derivative of (4.34) with respect to k_2 gives

$$\frac{\partial g_i}{\partial k_2} = \frac{B_{mi}\pi_i e^{-\pi_i t}}{(k_2 + \alpha_1 - \pi_i)^2} \quad (4.40)$$

92

Using (4.40), we obtain:

$$
\begin{aligned}
\frac{\partial f}{\partial k_2} &= \frac{B_{mi}\pi_i e^{-\pi_i t_d}}{(k_2 + \alpha_1 - \pi_i)^2} - \frac{B_{mi}\pi_i e^{-\pi_i t_1}}{(k_2 + \alpha_1 - \pi_i)^2} e^{-(k_2+\alpha_1)(t_d-t_1)} \\
&\quad + \left[\frac{B_{mi}}{(k_2 + \alpha_1 - \pi_i)}(k_2 + \alpha_1 - \pi_i - (k_2 + \alpha_1)e^{-\pi_i t_1}) \right](t_d - t_1)e^{-(k_2+\alpha_1)(t_d-t_1)} \\
&= \frac{B_{mi}e^{-(k_2+\alpha_1)(t_d-t_1)}}{(k_2 + \alpha_1 - \pi_i)^2} \Big[\pi_i e^{(k_2+\alpha_1-\pi_i)(t_d-t_1)} e^{-\pi_i t_1} - \pi_i e^{-\pi_i t_1} + (k_2 + \alpha_1 - \pi_i)^2(t_d - t_1) \\
&\quad -(k_2 + \alpha_1 - \pi_i)(k_2 + \alpha_1)(t_d - t_1)e^{-\pi_i t_1} + \pi_i e^{-\pi_i t_1}(k_2 + \alpha_1 - \pi_i)(t_d - t_1) \\
&\quad -\pi_i e^{-\pi_i t_1}(k_2 + \alpha_1 - \pi_i)(t_d - t_1) \Big] \\
&= \frac{B_{mi}e^{-(k_2+\alpha_1)(t_d-t_1)}}{(k_2 + \alpha_1 - \pi_i)^2} \Big[\pi_i e^{-\pi_i t_1} \left(e^{(k_2+\alpha_1-\pi_i)(t_d-t_1)} - (k_2 + \alpha_1 - \pi_i)(t_d - t_1) - 1 \right) + \\
&\quad + (k_2 + \alpha_1 - \pi_i)^2(t_d - t_1)(1 - e^{-\pi_i t_1}) \Big]
\end{aligned}
$$

As $e^m - m - 1 > 0$, it follows that $\frac{\partial f}{\partial k_2} > 0$ and that f increases monotonically as k_2 increases. Moreover :

$$
\frac{\partial z}{\partial k_2} = -(\bar{A} - |\Phi|)(t_d - t_1)e^{-(k_2+\alpha_1)(t_d-t_1)}
$$

As $|A| \leq \bar{A}$ (as an upper bound) and $|A| \geq |\Phi|$ (from A's definition), we have $\bar{A} \geq |\Phi|$. All the above leads to the conclusion that $\frac{\partial z}{\partial k_2} < 0$ and that z decreases monotonically as k_2 increases. Looking into (4.39) with the above results we note that as k_2 increases, t_d^+ decreases.

$\square$

4.4 Simulation results

In this section, extensive simulation results will be given to show potentialities and possible limitations of the proposed methodology. Specifically, a simple example is given just to emphasize some of the key aspects of the techique.

Consider the nominal system with $m = 1$, $K = 1$, $\sigma_0 = 2$, $\sigma_1 = \sqrt{2}$, $\sigma_2 = 0.4$, $f_c = 1$, $f_s = 1.5$, and $v_s = 0.001$. To implement the on-line approximator we have

employed HONNs, with sigmoid activation function $s(x) = \frac{m}{1+e^{-l(x-c)}} + \lambda$. Specifically, for the term $\hat{w}_2^\top S_2(|x_2|)$ we have chosen a 5th-order HONN with $(m, l, c, \lambda) = (0.8, -4, 2.119, -1.5)$, while for $\hat{w}_1^\top S_1(x_2, |x_2|)$ a 2nd-order HONN with $(m, l, c, \lambda) = (1.41, -10.0225, 0.5974, -2.11)$. To highlight the fault detectability issue, we first simulated the system with a fault of the form $\Delta F_1 = 20 + e^{10x_2}$ occurring at $T = 60$ sec (alteration in friction parameters). Then we simulated it for the type of fault, $\Delta F_2 = -1$ occurring at $T = 60$ sec which represents the spring's break. In both cases the design parameters were $k_1 = 100$ and $k_2 = 200$. The input u was $3\sin(0.2t)$. The results for faults ΔF_1 and ΔF_2 are depicted in Fig. 4.1 and Fig. 4.2, respectively. The detection time in which $|\tilde{\xi}_2| \geq \rho$, where ρ is the threshold defined in (4.27), for the first one was $t_d = 0.0076$ sec, while for the second one, $t_d = 0.0209$ sec. The subplots (4.1c-4.2c) depict the detectability condition ((4.32) or $|\int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)} B_i d\tau| - \rho - \rho e^{-(k_2+\alpha_1)(t_2-t_1)} - |\int_{T+t_1}^{T+t_2} e^{-(k_2+\alpha_1)(T+t_2-\tau)} A d\tau|$, and confirm the occurrence of the faults when becomes greater than zero. Fig. 4.3 shows the decreasing behavior of the detection time as a function of k_2.

4.5   Summary

In this chapter, we have presented an approach to detect faults in mechanical systems with friction that perform linear motion. The friction is modeled with the aid of the dynamic LuGre model. However all system nonlinearities and critical parameters are assumed unknown. Moreover, the frictional internal state is not available for measurement. The main contributions of this work are: 1) the development of an online neural network approximator for mechanical systems with friction that does not require full state measurement; and 2) the derivation of fault detectability conditions and the upper bounds of the detection time. Simulation results clarify and verify the theoretical analysis. In the following chapter, the DAMADICS benchmark problem is defined where the methodology developed here is applied yielding results that
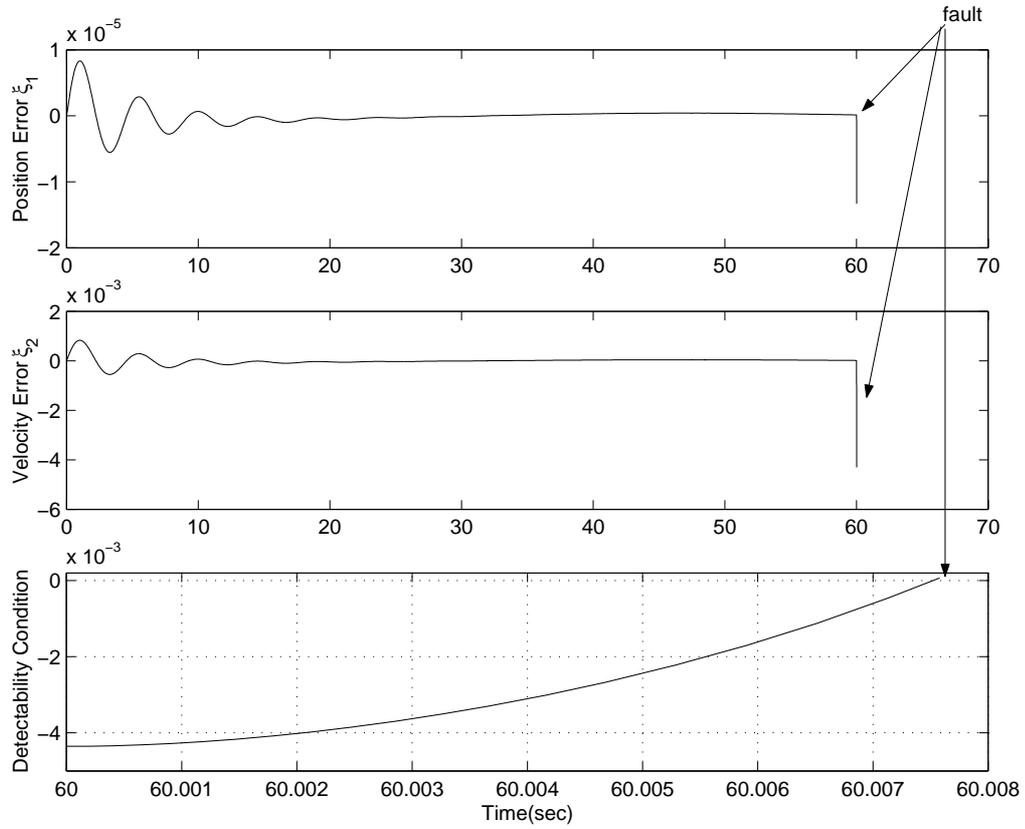
Figure 4.1: Behaviors of the: (a) position error $\tilde{\xi}_1 = x_1 - \hat{x}_1$; (b) velocity error $\tilde{\xi}_2 = x_2 - \hat{x}_2$; (c) detectability Condition
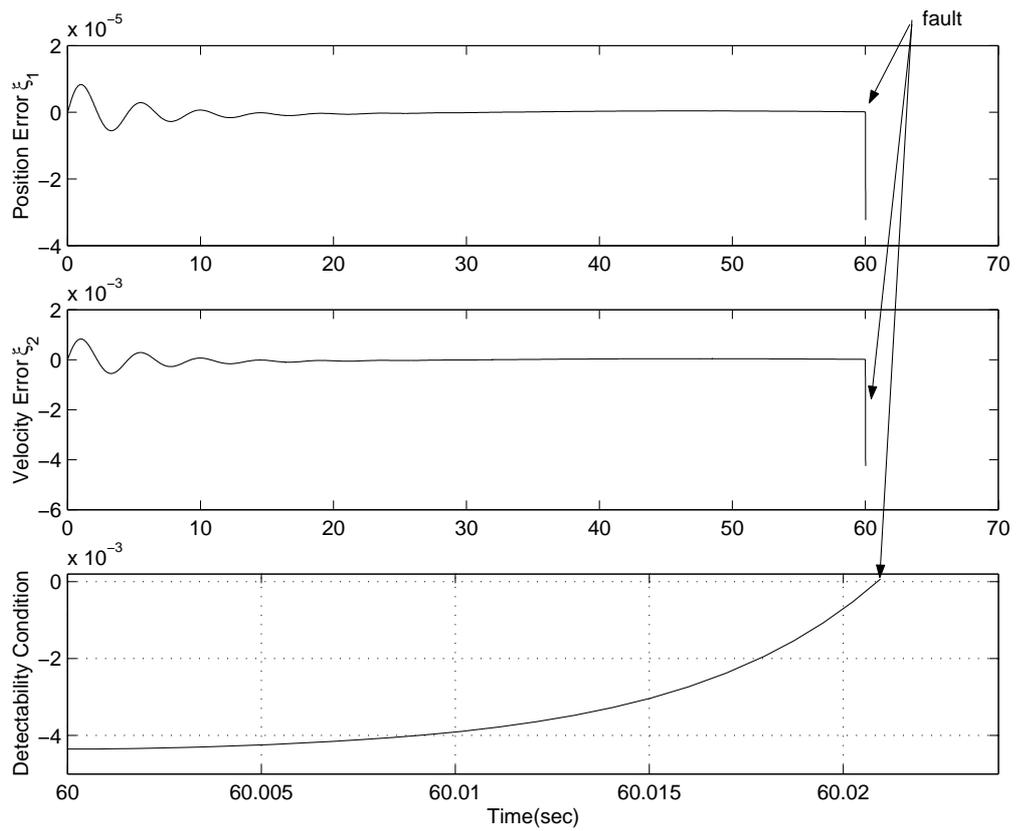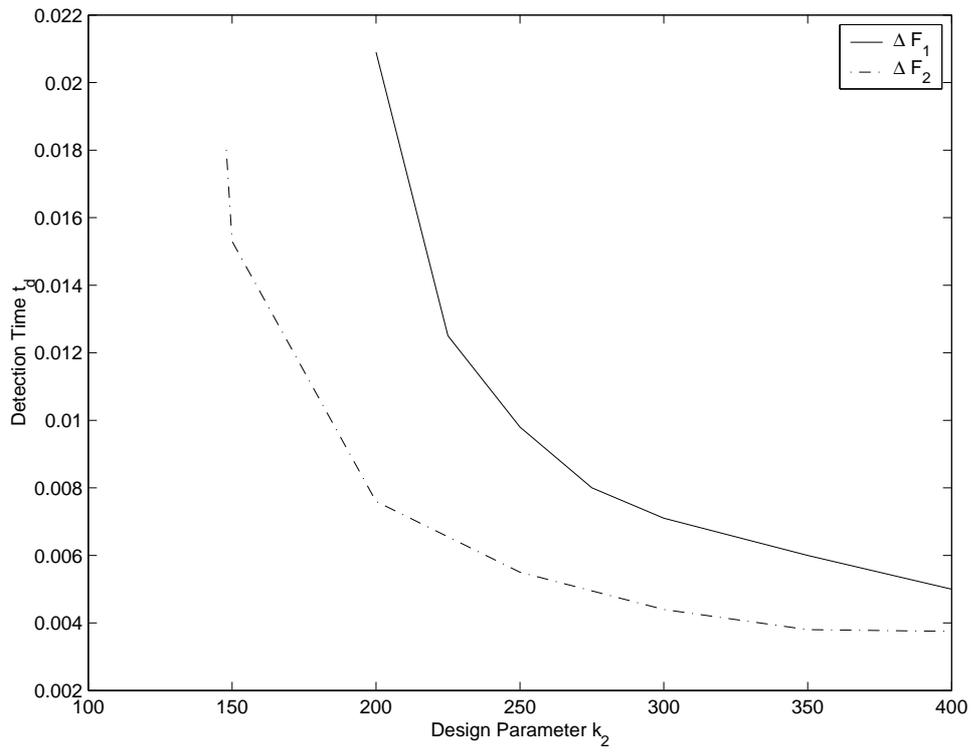
clarify and verify, additionally, its reliability.

Figure 4.2: Behaviors of the: (a) position error $\tilde{\xi}_1 = x_1 - \hat{x}_1$; (b) velocity error $\tilde{\xi}_2 = x_2 - \hat{x}_2$; (c) detectability Condition

Figure 4.3: Dependence of detection time t_d on the value of parameter k_2

Chapter 5

Fault Detection in Damadics benchmark problem

In this chapter we concentrate on detecting faults giving emphasis to the DAMADICS[1] actuator[2] benchmark problem applying the methodology developed in Chapter 4. In the framework of the DAMADICS research network funded by the European Union, a benchmark model was developed to approximate the behavior of the evaporation stage of a sugar factory in Lublin (Poland). Actuators under consideration consist of a control valve, a pneumatic linear servomotor and a positioner. In such a kind of electromechanical systems, the presence of friction phenomena is unavoidable and significantly increases the complexity of the FD problem.

[2]Actuator or a final control element is a physical device, structure or assembly of devices acting on controlled process

5.1   Plant description

The plant under concern is the sugar factory Cukrownia Lublin S.A located in Lublin (Poland). Specifically, we consider the evaporation process where the main task is to thicken the beet juice coming from the cleaning and filtering stages, at the minimum heat-energy consumption. The first three sections work with natural juice circulation and the last two work with juice circulation forced by pumps. We focus on the first section, consisting of one evaporator and containing an important actuator, located on the inflow of thin juice and controlling its level in the first stage of evaporation station.
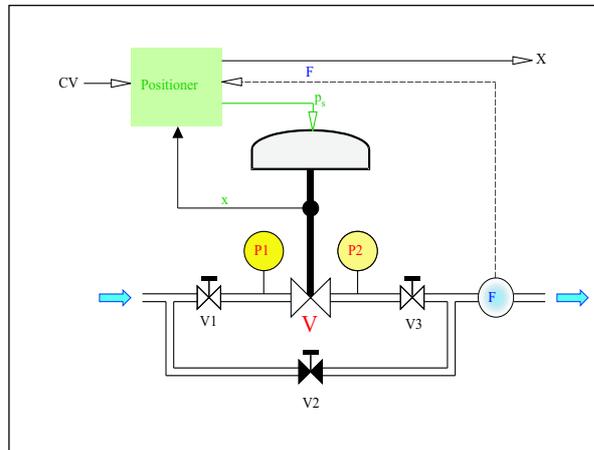

Figure 5.1: A control valve-pneumatic servomotor-positioner device.

As shown in Fig. 5.1, the actuator is made of three main components [55]:

- Control valve driven by a servomotor, which is used to prevent, to allow and/or to limit the flow of fluids.

- Spring-and-diaphragm pneumatic servomotor; this is a compressible fluid powered device where the fluid acts upon the flexible diaphragm thus providing linear motion of the servomotor stem.

Symbols	Meaning		Symbols	Meaning
k_s	Spring constant		F_{fV}	Viscosity friction force
k_d	Diaphragm constant		F_{fC}	Coulomb friction force
p_s	Air pressure in chamber		F_{vc}	Vena-contracta force
F_n	Normal packing force		F_{dA}	d'Alambert force
F_p	Active force		x	Rod's displacement
F_g	Gravity force		x_0	Initial spring compression
F_s	Spring compression force		m	mass of rod, valve, diaphragm

Table 5.1: Explanation of the symbols of the pneumatic servomotor and its physical layout.

- Positioner; this device is used to eliminate control-valve stem miss-positions due to external or internal sources such as friction, hydrodynamic forces, etc..

Fig. 5.2 shows a more detailed overview of the servomotor as well as its physical layout; the effects (forces) of the other two components are emphasized (the meaning of the symbols is straightforward and is presented in Table 5.1).
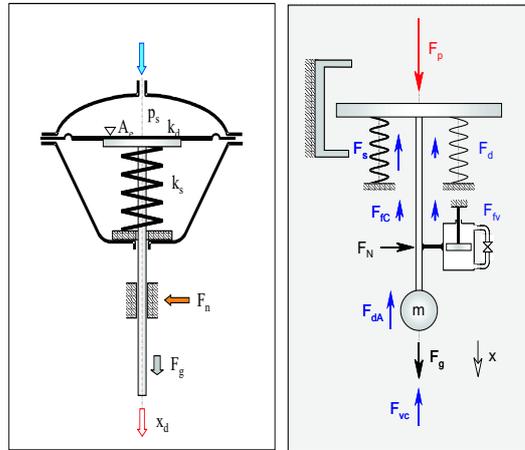

Figure 5.2: The pneumatic servomotor and its physical layout.

A rather detailed dynamic model of the above evaporation process (and of the actuator as well) has been developed and validated in the context of the DAMADICS research training network. The unavoidable friction effects are modelled by means of

suitable hysteresis functions. In this work instead, the frictional effects are described by the already mentioned dynamic model, the LuGre model. The reason to use this friction model is that it is able to capture important phenomena such as presliding displacement, frictional lag, stick-slip motion, etc.. Another important reason is that, in the considered actuator, the motion corresponds to a low-velocity motion. In such a case, the friction nonlinearities dominate and the LuGre model is very suitable to characterize these nonlinear effects.

5.2 Problem formulation for DAMADICS case

The linear motion provided by the servomotor device, the use of LuGre model as well as the fault definition given by the DAMADICS benchmark motivated us to apply our developed methodology. It is important to clarify that the above-described dynamic model for a mechanical system with friction phenomena in both nominal and faulty modes of operation has a different structure with respect to the DAMADICS model. However, the complexity of the DAMADICS model rules out the possibility of using it in the framework of a nonlinear model-based FD algorithm. Therefore, the key idea is to determine a suitable LuGre model to make its behavior very similar to the DAMADICS one from an input–output perspective. This will allow us to use the LuGre model to design a model-based FD scheme as it is described in the previous sections. In Fig. 5.3, this intuitive idea is shown in a schematic way.[3]

In this respect using the theoretical results, the approximator's output will serve as the residual signal for fault detection. Owing to the convergence analysis presented

---

[3]The use of the LuGre model needs the velocity measurement which is not available in the DAMADICS actuator case. However, the velocity can be easily estimated using the position measurements by means of a suitably designed Kalman filter.
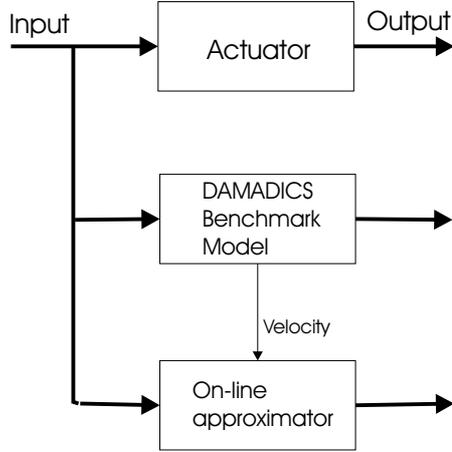
Figure 5.3: Architecture of the adaptive on–line approximation scheme.

before, it follows that Φ can be used to define the threshold function ρ as:

$$\rho = \frac{|\Phi|}{k_2} \tag{5.1}$$

Choosing now as a residual signal $\tilde{\xi}_2$ with its correspondent threshold (5.1), we can say that a fault will be detected when $\tilde{\xi}_2 \geq \rho$.

5.3   Damadics Simulation Results

In this part we present the simulation results regarding actuator faults introduced in friction and servomotor's spring. Relative to the friction fault an increasing of valve or bushing friction is considered. Mechanical wear, air pollution, corrosion products and sedimentation consist the reasons of existence and the physical interpretation of the fault related to friction. On the other hand relative to the servomotor's spring fault, the harsh environment causes fatigue or corrosion of spring material. The results were taken according to the scheme that is depicted in Fig. 5.4.

P_1 and P_2 represent the pressure before and after the control valve and were set to be $3.5 \cdot 10^6$Pa and $2.6 \cdot 10^6$Pa respectively. T represent the water temperature and was 20^oC. CV is the control value that takes values in $[0, 1]$. A value of "1"
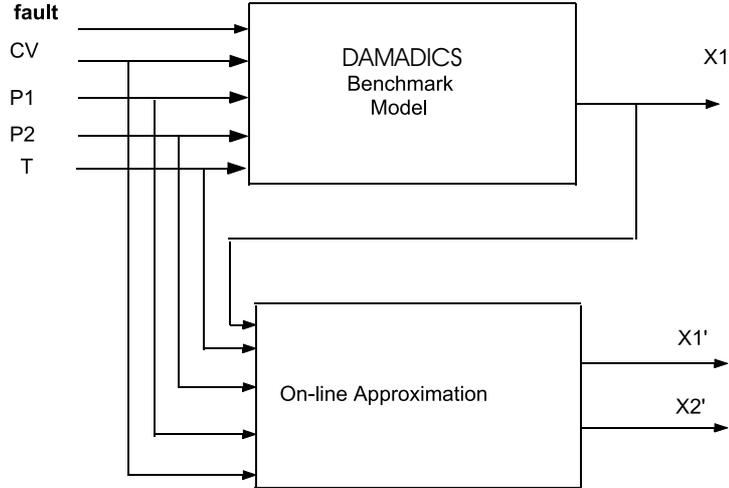
Figure 5.4: Architecture used for DAMADICS simulation trials.

expresses that the valve is closed where a value of "0" a fully-opened valve. The output of the benchmark model $X1$, represents the rod's displacements. To implement the on-line approximator we have employed High Order Neural Networks (HONNs), with sigmoid activation function of the form $s(x) = \frac{m}{1+e^{-l(x-c)}} + \lambda$. Specifically for the term $\hat{w}_2^\top S_2(|x_2|)$ we have chosen a 5th-order HONN with $(m, l, c, \lambda) = (0.8, -4, 2.119, -1.5)$, while for $\hat{w}_1^\top S_1(x_2, |x_2|)$ a 2nd-order HONN with $(m, l, c, \lambda) = (1.41, -10.0225, 0.5974, -2.11)$. The design constants k_1 and k_2 were set to be 100 and 400 respectively. The outputs $X1'$ and $X2'$ of the on-line approximator represent the estimated position and velocity respectively. As simulations have been carried out in a noise-free environment, the velocity was estimated by introducing a high-pass filter. According also to benchmark definition, the faults are standardized to the range of $[-1\ 1]$. The limiting values "-1" and "1" corresponds to some pre-defined states or physical values (Δf_{min}, Δf_{max}). Fault notations are given in Table 5.2.

The type of faults can be either abrupt or incipient. More specifically, the fault concerning friction is an incipient one. The fault that we simulated occurs at $t = 70$

103

Friction fault	Servomotor spring fault
−1 - no friction	−1 - spring's perforation
0 - unchanged friction	0 - no fault
1 - advanced friction	1 - spring's tightness

Table 5.2: Fault specifications.

sec and takes its final value "1" after 20 sec. A detection decision (0-no fault, 1-fault) is being made when $|\tilde{\xi}_2| \geq \rho$ for more than one sample time. The simulations results are depicted in Fig. 5.5.

The conclusions that can be drawn from Figs. 5.5(a)-(b) are that the adaptive scheme is able to learn on line the behavior of the model with very small errors. In Fig. 5.5(c) a parallel graph of $|\tilde{\xi}_2|$ and of the corresponding threshold ρ is plotted. As it mentioned before, a fault decision is taken when $|\tilde{\xi}_2| \geq \rho$ for more than one sample time. Specifically:

$$|\tilde{\xi}_2(t)| \quad \geq \quad \rho(t)$$

$$AND$$

$$|\tilde{\xi}_2(t + \Delta t)| \quad \geq \quad \rho(t + \Delta t)$$

where t is the time instant at which $|\tilde{\xi}_2| \geq \rho$ and Δt is the sampling step. This is the reason why no fault indication is turned on before the actual occurrence of the fault (see Fig. 5.5(e)), despite some spikes occurring before the time of fault occurrence (see Fig. 5.5(c)).

As can be noticed from Fig. 5.5(e), the fault is detected at $t = 81.12$sec. The fault strength on this time–instant is of about 50% of its final value (Fig. 5.5(d)), a characteristic which can prevent on time the overall system from serious damages. Similar comments can be made when we simulate the system with the servomotor spring fault (see Fig. 5.6), which, according to the benchmark definition, is an abrupt fault. In this case, the fault is detected at $t = 70.005$sec.
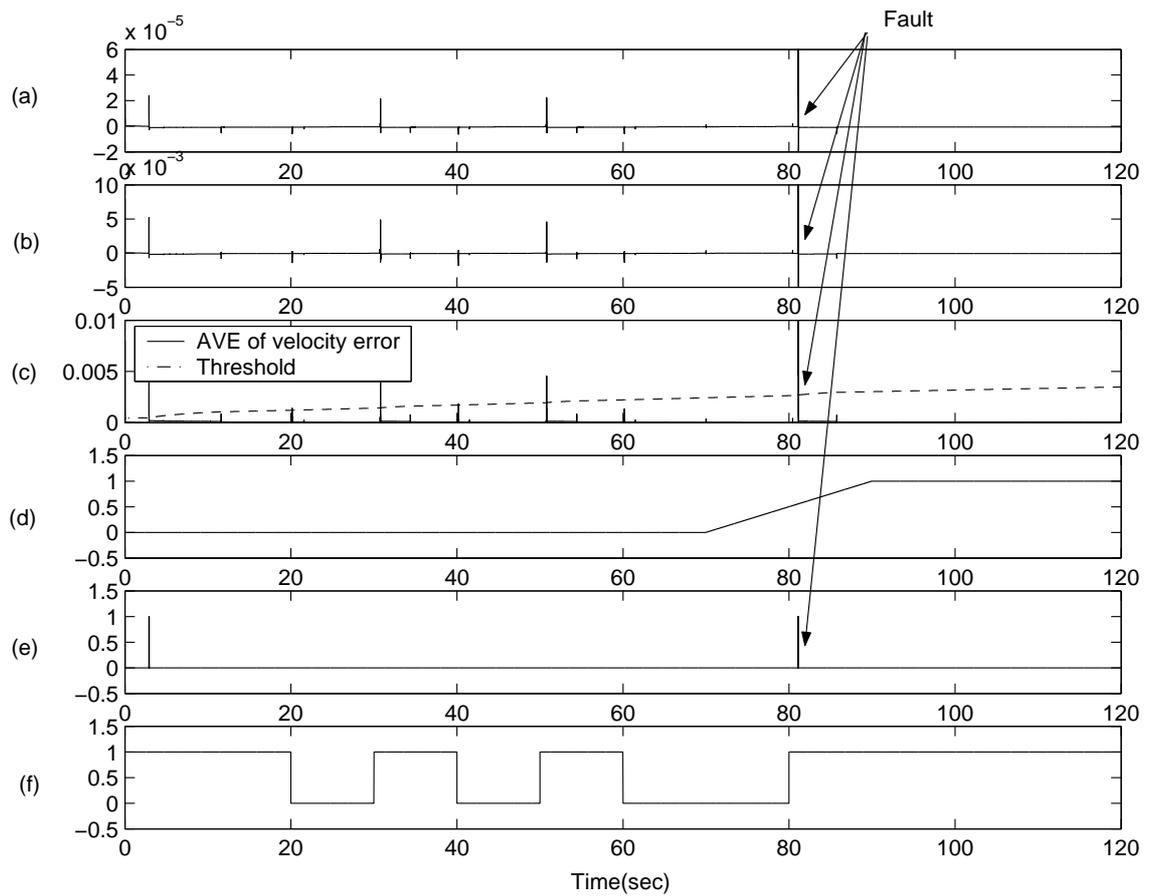
104

Figure 5.5: Behaviors of (a) position error $\tilde{\xi}_1$. (b) velocity error $\tilde{\xi}_2$. (c) absolute value of velocity error, $|\tilde{\xi}_2|$ threshold ρ. (d) A fault evolution. (e) detection decision. (f) control Value (CV).
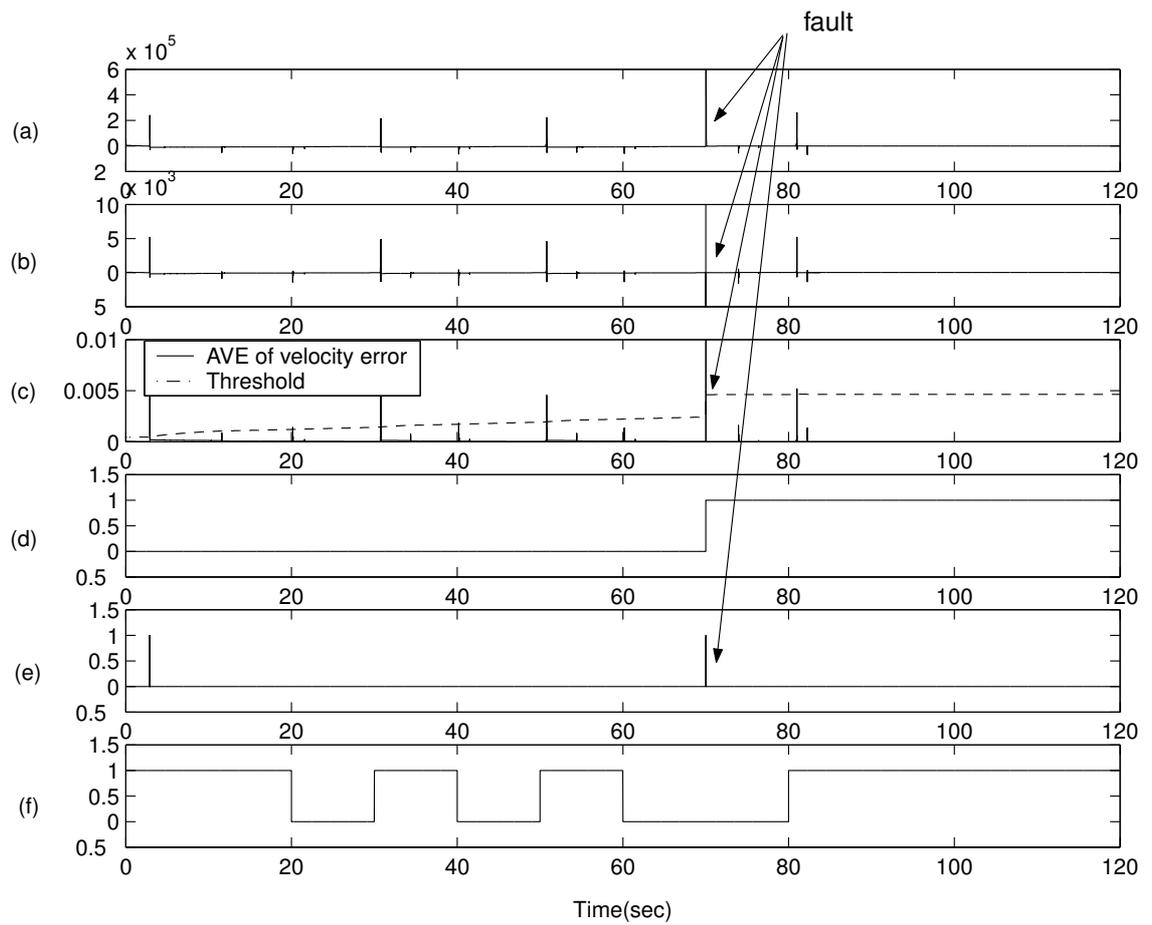
Figure 5.6: Behaviors of (a) position error $\tilde{\xi}_1$. (b) velocity error $\tilde{\xi}_2$. (c) absolute value of velocity error, $|\tilde{\xi}_2|$ threshold ρ. (d) A fault evolution. (e) detection decision. (f) control Value (CV).

106

Bibliography

[1] Arkan Kayihan, Francis J. Doule III, "Friction compensation for a process control valve," *Control enginnering practice*, 8, pp.799-812, 2000.

[2] B. Armstrong, "Dynamics for robot control: friction modeling and ensuring excitation during parameter identification," Ph.D thesis, Dept. of Electrical Engineering, Stanford University, Stanford Computer Science Memo STAN-CS-88-1205, 1988.

[3] B. Armstrong-Hélouvry, *Control of Machines with Friction.* Kluwer Academic Publishers, Norwell, MA, 1991.

[4] B. Armstrong-Hélouvry, "A perturbation analysis of stick-slip. In R. A. Ibrahim and A. Soom (Eds)," *Friction-Induced Vibration, Chatter, Squeal, and Chaos, Proc. ASME Winter Annual Meeting, Anaheim,* DE-Vol. 49, ASME, NY, pp.41-48.

[5] B. Armstrong-Hélouvry, B. and P. Dupont, "Friction modeling for controls, and Compensations Techinques for Servos with Friction," in *Proc. American Control Conference*, pp.1905-1915, 1993.

[6] B.Armstrong-Hélouvry, P. Dupont and Canudas de Wit , "A survey of models, analysis tools and compensation methods for the control of machines with friction," *Automatica*, vol. 30, No. 7, pp.1083-1138, 1994.

[7] Canudas de Wit, K.J. Åström, and K. Braun, "Adaptive friction compensation in DC motor drives", in *IEEE Conference on Robotics and Automation,*, 3:1556-1561, December, 1986.

[8] Canudas de Wit, P. Noel, A. Auban and B. Brogliato, "Adaptive Friction Compensation in Robot Manipulators: Low Velocities," International Journal of Robot Reasearch, vol. 10, No. 3, pp.189-199, 1991.

[9] Canudas de Wit, H. Olsson, K.J. Åström, and P. Lischinsky, "A New Model for Control of Systems with Friction," *IEEE Trans. on Automatic Control*, vol. 40, No. 3, pp.419-425, March 1995.

[10] Canudas de Wit, S.S. Ge, "Adaptive Friction Compensation for Systems with Generalized Velocity/Position Friction Dependency", in *Proc. 36th Conf Decis Contr,* pp.2465-2470, December, 1997.

[11] Chua L.O and Stromsmoe, "Mathematical model for dynamic hysteresis loops," *Int.J.Eng.Sci.,* vol. 19, pp. 435-450, 1971

[12] Chua L.O and Bass S.C, "A generalized Hysteresis model," *Trans. on Circuit Theory,* vol. CT-19, No. 1, pp. 36-48, January 1972.

[13] Chun-Yi Su, Y. Tan, Y. Stepanenko, "Adaptive Control of a Class of Nonlinear Systems Preceded by an Unknown Backlash-Like Hysterisis", in *Proc. 39th Conf Decis Contr,* 2000.

[14] Cincinatti Milacron, Revised Stick-Slip Test Procedure, 1986.

[15] Dahl, P.R "A solid friction model", TOR-158(3107-18), The Aerospace Corporation, El Segundo, CA, 1968.

[16] Dahl, P.R "Solid friction damping of mechanical vibrations," *AAIA J.,* **14**(12), pp.1675-1682, 1976.

[17] Dahl, P.R "Measurement of sloid friction parameters of ball bearings," *Proc of 6th Annual Symp on Incremental Motion, Control Systems and Devices*, University of Illinois, ILO, 1977.

[18] J.R. Rice and A.L. Ruina, "Stability of steady frictional slipping," *J. Applied Mechanics*, vol. 50, No.2, 1983

[19] G. Tao and P.V. Kokotoviç, "Continuous-time Adaptive control of systems with unknown backlash," *IEEE Trans. on Automatic Control*, vol. 40, pp.1083-1087, 1995a.

[20] G. Tao and P.V. Kokotoviç, "Adaptive control of plants with unknown hystereses," *IEEE Trans. on Automatic Control*, vol. 40, pp.200-212, 1995b.

[21] G. Tao and P.V. Kokotoviç, *Adaptice Control of Systems with Actuator and Sensor Nonlinearities*, Wiley Interscience, NY, 1996.

[22] J. Swevers, Farid Al-Bender, Chris G.Ganseman and Tutuko Prajogo, "An Integrated friction model structure with improved presliding behavior for accurate friction compensation,",*IEEE Trans. on Automatic Control*, vol. 45, No. 4, pp. 675-686, April 2000.

[23] Mayergoyz I.D., *Mathematical models of Hysteresis.* New York: Springer-Verlag, 1991

[24] Macki J.W., Nistri P. and Zecca P., "Mathematical models for hysteresis," *SIAM Review*, vol. 35, pp.94-123, 1993.

[25] S. Kato, K. Yamaguchi and T. Matsubayashi, "Some considerations of characteristics of static friction of machine tool slideway," *J. of Lubrication Technology*, **94**(3), pp.234-247, 1972.

[26] Gertler J.J., *Fault Detection and Diagnosis in Engineering Systems* New York: Marcel Dekker, 1998

[27] Chen J., Patton R.J., *Robust Model-based Fault Diagnosis for Dynamic Systems* Massachusetts, USA: Kluwer Academic Publishers, 1999

[28] P.M. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy - a survey and some new results," *Automatica,* vol. 26, pp.459-474, 1990.

[29] J.J. Gertler, "Fault detection and isolation using parity relations," *Control Eng. Practice*, vol. 5, No. 5, pp.653-661, 1997.

[30] Gertler J.J., "Analytical redundancy methods in fault detection and isolation," *Proc. of the IFAC/IMACS Symposium SAFEPROCESS*, pp.9-21.

[31] Magni J.F. and Ph. Mouyon, "On the residual generation by observer and parity space approaches," *IEEE Trans. on Automatic Control*, vol. 39, pp.441-447, 1994

[32] Ding S.X, E.L. Ding and T.Jeinsch, "An approach to analysis and design of observer and parity relation based FDI systems," *Proc. The 14th IFAC World Congr. Beiijing*, 1999

[33] Zhuang Z. and P.M Frank, "A fault detection scheme based on stochastic qualitative modeling," *Proc. The 14th IFAC World Congr. Beiijing*, 1999

[34] Kuipers B. "Qualitative simulation," *Artificial Intelligence*, vol 66, No. 29, pp.289-338, 1986

[35] Kay H. and B. Kuipers, "Numerical behavior envelopes for qualitative models," *Proceedings of the 7th National Conference on Artificial Intelligence,* pp.606-613, 1993.

[36] Bonarini A. and G. Bontempi, "A qualitative simulation approach for fuzzy dynamical models", *ACM transactions on Modeling and Computer Simulation*, vol 4, No. 4, pp.285-313, 1994.

[37] Lunze J. "Qualitative modeling of linear dynamical systems with quantized state measurements", *Automatica*, vol 30, No. 3, pp.417-431, 1994.

[38] Chen J., Lopez-Toribio C.J and Patton R.J., "Non-linear dynamic systems fault detection and isolation using fuzzy observers", *Proc. Instn. Mech. Engrs.* **213, Part 1**, 1999.

[39] Marcu T., Matcovschi M.H. and Frank P.M., "Neural observer-based approach to fault detection and isolation of a three-tank system", *ECC'99, Karlsruhe*, 1999.

[40] Marcu T., Matcovschi M.H. and Frank P.M., "Neural observer-based approach to fault tolerant control of a three-tank system", *ECC'99, Karlsruhe*, 1999.

[41] Calado J.M.F. and J.M.G. Sa da Costa, "On-line fault detection and diagnosis based on a coupled system", *ECC'99, Karlsruhe*, 1999.

[42] N.E. Cotter, "The Stone-Weierstrass theorem and its applications to neural networks", *IEEE Trans. on Neural Networks*, vol.1, pp.290-295, 1990.

[43] G.Cybenko, "Approximations by superpositions of a sigmoidal function", *Mathematics of Control, Signals, and Systems*, vol. 2 pp.303-314, 1989.

[44] *Neuro-Control Systems: Theory and Applications.* M.M. Gupta and D.H. Rao (Eds.), New York, IEEE Press, 1994.

[45] P.A. Ioannou and J. Sun. *Robust Adaptive Control.* Englewood Cliffs, NJ:Prentice Hall, 1995.

[46] T. Poggio and F. Girosi, "Regularization algorithms for learning that are equivalent to multilayer networks", *Science*, vol. 247, pp.978-982, 1990.

[47] G.A. Rovithakis and M.A. Christodoulou. *Adaptive Control with Recurrent High-Order Neural Networks.* Springer-Verlag, London, 2000

[48] *Handbook of Intelligent Control: Neural, Fuzzy and Adaptive Approaches.* D.A. White and D.A. Sofge (Eds.). New York, IEEE Press, 1994

[49] A. R. Barron, "Universal approximation bounds for superpositions of a sigmoidal function," *IEEE Trans. on Information Theory*, vol. 39, pp. 930–945, 1993.

[50] R. Zoppoli, M. Sanguineti, and T. Parisini, "Approximating networks and extended Ritz method for the solution of functional optimization problems," *Journal of Optimization Theory and Applications*, vol. 112, n. 2, pp. 403–439, 2002.

[51] M. M. Polycarpou and A. B. Trunov, "Learning Approach to Nonlinear Fault Diagnosis: Detectability Analysis," *IEEE Trans. on Automatic Control,* Vol. 45, No. 4, pp.806-812, 2000.

[52] M.A. Demetriou and M.M. Polycarpou, "Incipient fault diagnosis of dynamical systems using on-line approximators," *IEEE Trans. on Automat. Contr.,* vol. 43, pp.1612-1617, 1998.

[53] M.M. Polycarpou M.A. Helmicki, "Automated fault detection and accomodation: A learning approach," *IEEE Trans. Syst. Man, Cybern.,* vol. 25, pp.1447-1458, 1995.

[54] E. Alcorta Garcia and P.M. Frank, "Deterministic nonlinear observer-based approaches to fault diagnosis: a survey," *Control Eng. Practice*, vol. 5, No. 5, pp.663-670, 1997.

[55] M.Z. Bartys, "Specification of actuators intented to use for benchmark defini-
tion", Technical Report, Poland, 2002.

Vita

...

Permanent Address: Solonos 101 str.

P.O. 10678 Athens

GREECE

This thesis was typeset with $\text{\LaTeX}\,2_\varepsilon$[4] by the author.

---

[4]$\text{\LaTeX}\,2_\varepsilon$ is an extension of $\text{\LaTeX}$. $\text{\LaTeX}$ is a collection of macros for $\text{\TeX}$. $\text{\TeX}$ is a trademark of the American Mathematical Society. The macros used in formatting this thesis were written by Dinesh Das, Department of Computer Sciences, The University of Texas at Austin, and extended by Bert Kay and James A. Bednar.