Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-11 (2016) 547–554

A comprehensive observer-based fault isolation method with application to a hydraulic power train

Sebastian Pröll * **Fabian Jarmolowitz** * **Jan Lunze** **

* *Robert Bosch GmbH, Germany (e-mail: {sebastian.proell, fabian.jarmolowitz}@de.bosch.com).*
** *Ruhr-University Bochum, Germany, (e-mail: jan.lunze@ruhr-uni-bochum.de)*

Abstract: This paper concerns observer-based fault isolation. As a new contribution it combines structural diagnosability analysis with observer-based residual generation. Structural analysis techniques are used to find diagnosable subsystems within the system model from each of which an observer can be built. The resulting bank of observers provides a vector of residuals. Evaluation of this residual vector is realized by structural methods which makes fault isolation possible. The result is a comprehensive way to build a fault diagnoser for a class of nonlinear systems. The proposed diagnosis method is evaluated at an automotive application, which is a hydraulic power train.

Keywords: Observer-based residual generation, fault isolation, structural analysis, hydraulic power train

1. INTRODUCTION

Automotive systems exhibit an ever increasing level of automation and complexity. A malfunction of such systems may cause significant damage, pollution or danger to humans. Supervision is mandatory in order to maintain safety and reliability, and to fulfill legal requirements. The task of diagnostic algorithms is to detect and isolate faults that may affect the system. Fault detection is concerned with verifying whether or not a system is faulty. Fault isolation tries to find the faulty component in the system.

Automotive systems consist of various components, including actuators, sensors, and control units. Due to the size of those systems it is no longer advisable to design diagnostic functions that are based on heuristic analysis or expert knowledge. Instead, systematic methods have to be developed in order

- to analyze whether or not faults are diagnosable, i.e., detectable or isolable, respectively, and
- to build diagnostic algorithms.

The present paper provides a comprehensive fault isolation method which tackles the two problems and which is based on structural analysis on the one hand and observer-based diagnosis on the other hand.

Structural analysis has become popular in model-based diagnosis to analyze diagnosability (Blanke et al. (2016)) and has been proved to be effective for automotive systems, see, e.g., Svärd and Nyberg (2010). This approach takes a mathematical model of the system and abstracts from the analytical expressions to the system structure, where only the connections between the variables and the model equations are considered. A graph-theoretic tool for bipartite graphs is used to find over-determined subsets of equations (model $\mathcal{C}^+$ in Fig. 1) within the system model. These include redundancies and make fault diagnosis possible. By using the information which model equations are affected by which fault, diagnosability analysis is performed and a residual evaluation logic is provided.

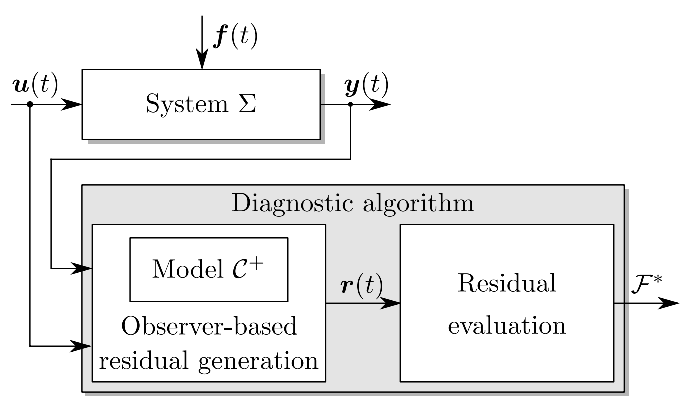

Fig. 1. Diagnosis of system Σ

Diagnostic algorithms use the known input and output signals of a system to compute residuals (signal $\boldsymbol{r}(t)$ in Fig. 1) that indicate whether the system behavior is consistent with the model of the non-faulty system. By evaluating these residuals appropriately, the fault isolation task can be accomplished.

Literature. Residual generation in the structural analysis community is often based on computing analytical redundancy relations (ARRs) which are functions of the input and output signals and their derivatives (Staroswiecki and Comtet-Varga (2001); Düştegör et al. (2006)). Noisy measurements pose a serious problem for this approach.

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.08.080

Moreover, it is not clear whether it is possible to compute a residual from an over-determined set of equations. Especially for dynamic systems this becomes relevant, and different ways how to deal with this problem have been proposed (Åslund et al. (2011); Frisk et al. (2012)).

Observer-based methods by contrast are well established to generate residuals for dynamic systems (Isermann (2005)). Their strength relies on the robustness of the residual generators against measurement noise. The goal of the present paper is to exploit the strengths of both the structural analysis and the observer-based residual generators. That is, it starts with a state space model of the system which shall be diagnosed, and finds over-determined subsets of equations by a structural analysis. From these subsets, those sets are selected which define a state space model and make an observer-based residual generation possible. The fault isolation step, in which residuals are evaluated, is again realized by structural techniques. Thus, a systematic approach towards a fault diagnosis algorithm is provided in this paper. The fault isolation procedure is evaluated at a simulation model of an automotive application, which is a hydraulic power train.

Structure of the paper. In Section 2 basic definitions from structural analysis are revisited and it is explained how fault isolation by means of structural methods can be established. Section 3 provides the novel result which is a fault diagnosis approach that combines structural analysis and observer-based residual generators for fault isolation. A procedure is presented in Section 4 that explains step by step how to get to the fault diagnoser. The presented fault diagnosis method is evaluated at a hydraulic power train in Section 5.

2. STRUCTURAL ANALYSIS

This section introduces the model class considered in this paper and provides the structural analysis concepts that are used later to build a fault diagnoser for fault isolation. It is explained how the structural model is obtained from a given analytical model and the notion of structurally over-determined sets is revisited.

2.1 Model class and structural model

Consider the nonlinear system in state space representation

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}(t)), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \\ \boldsymbol{y}(t) &= \boldsymbol{h}(\boldsymbol{x}(t)), \end{aligned} \tag{1}$$

with an input vector $\boldsymbol{u}(\cdot) \in \mathbb{R}^m$, a state vector $\boldsymbol{x}(\cdot) \in \mathbb{R}^n$, and an output vector $\boldsymbol{y}(\cdot) \in \mathbb{R}^p$. The output function $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^p$ is of the form

$$y_i(t) = h_i(x_i(t)), \quad i = 1, \ldots, p, \tag{2}$$

with invertible functions $h_i : \mathbb{R} \to \mathbb{R}$. That is, the ith measurement $y_i(t)$ only depends on the ith system state $x_i(t)$ for all $i = 1, \ldots, p$, and a direct mapping $x_i(t) = h_i^{-1}(y_i(t))$ is possible.

Subsets of equations within the model (1) are aimed to be found which themselves define a state space model. For this, methods from structural analysis are used which shall be introduced subsequently.

In structural analysis, the model (1) is interpreted as a set $\mathcal{C}$ of equations

$$\mathcal{C} : \begin{cases} e_1 : & \dot{x}_1(t) = g_1(\boldsymbol{x}(t), \boldsymbol{u}(t)), \\ & \vdots \\ e_n : & \dot{x}_n(t) = g_n(\boldsymbol{x}(t), \boldsymbol{u}(t)), \\ m_1 : & y_1(t) = h_1(x_1(t)), \\ & \vdots \\ m_p : & y_p(t) = h_p(x_p(t)), \\ d_i : & \dot{x}_i(t) = \dfrac{\mathrm{d}}{\mathrm{d}t} x_i(t), \quad i = 1, \ldots, n. \end{cases} \tag{3}$$

In this representation, e_i, $(i = 1, \ldots, n)$, and m_i, $(i = 1, \ldots, p)$, denote the names of the state and output equations, respectively. In structural analysis, variables are related to each other only via equations. Thus, the relationship between $x_i(t)$ and its derivative $\dot{x}_i(t)$ has to be made explicit via *differential constraints* d_i, $(i = 1, \ldots, n)$. The equations are summarized in the set

$$\mathcal{C} = \{d_1, \ldots, d_n, e_1, \ldots, e_n, m_1, \ldots, m_p\},$$

and the unknown variables in the set

$$\mathcal{X} = \{x_1, \ldots, x_n, \dot{x}_1, \ldots, \dot{x}_n\}.$$

For a set of equations C define

$$\mathrm{var}(C) := \{z \in \mathcal{X} \mid z \text{ appears in an equation } c \in C\}.$$

The structural model of (1) is given by means of a bipartite graph

$$\mathcal{G} = (\mathcal{C}, \mathcal{X}, \mathcal{E}), \tag{4}$$

which consists of the two vertex sets $\mathcal{C}$ and $\mathcal{X}$, and an edge set $\mathcal{E} \subseteq \mathcal{C} \times \mathcal{X}$, which is defined as follows:

$$(c_i, z_j) \in \mathcal{E} \quad \text{if} \quad z_j \in \mathrm{var}(\{c_i\}).$$

The goal in analyzing this graph is to find over-determined subsets of equations in (3) which contain more equations than unknown. These sets are candidates for residual generation.

2.2 Over-determined sets of equations

In order to find over-determined subsets $P \subseteq \mathcal{C}$, the structural model by means of the *incidence matrix* $\boldsymbol{M}$ is introduced. The rows of this matrix represent the equations of $\mathcal{C}$, the columns represent the unknowns of $\mathcal{X}$. Thus, the dimensions of $\boldsymbol{M}$ are $(2n + p) \times 2n$. $\boldsymbol{M}$ is defined as follows:

> For arbitrary $c_i \in \mathcal{C}$ and $z_j \in \mathcal{X}$, the corresponding entry in $\boldsymbol{M}$ is 1 if $(c_i, z_j) \in \mathcal{E}$, else it is 0.

By the same rule, the submatrices $\boldsymbol{G}$ and $\boldsymbol{H}$ can be defined which only consider the edges in $\{e_1, \ldots, e_n\} \times \{x_1, \ldots, x_n\}$ and $\{m_1, \ldots, m_p\} \times \{x_1, \ldots, x_n\}$, respectively. Thus, the incidence matrix of the model (3) has the form given in Fig. 2. By a permutation of rows and columns of $\boldsymbol{M}$, which is known as the Dulmage-Mendelsohn decomposition of $\boldsymbol{M}$ (Dulmage and Mendelsohn (1958); Pothen and Fan (1990)), the matrix can be brought into block triangular form and the decomposition yields a unique partition of the sets $\mathcal{C}$ and $\mathcal{X}$. More precisely, $\mathcal{C} = \mathcal{C}^0 \cup \mathcal{C}^+$ and $\mathcal{X} = \mathcal{X}^0 \cup \mathcal{X}^+$, and for the respective sets it holds that

$$|\mathcal{C}^0| = |\mathcal{X}^0|, \quad |\mathcal{C}^+| > |\mathcal{X}^+|.$$

$\mathcal{C}^+$ is called the *structurally over-determined part* of (3) (Blanke et al. (2016)). Within $\mathcal{C}^+$, further *proper structurally over-determined* (PSO) subsets $P \subseteq \mathcal{C}^+$ can be